

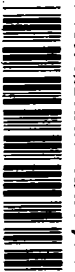
**NASA CONTRACTOR
REPORT**

NASA CR-944



NASA
CR
935
v.10
c.1

NASA CR-944



0060031

TECH LIBRARY KAFB, NM

LOAN COPY: RETURN TO
AFWL (WLIL-2)
KIRTLAND AFB, N MEX

**DYNAMIC STABILITY
OF SPACE VEHICLES**

Volume X - Exit Stability

by R. F. Ringland

Prepared by
GENERAL DYNAMICS CORPORATION
San Diego, Calif.
for George C. Marshall Space Flight Center



DYNAMIC STABILITY OF SPACE VEHICLES

Volume X - Exit Stability

By R. F. Ringland

Distribution of this report is provided in the interest of information exchange. Responsibility for the contents resides in the author or organization that prepared it.

Issued by Originator as GDC-BTD-67-029

Prepared under Contract No. NAS 8-11486 by
GENERAL DYNAMICS CORPORATION
San Diego, Calif.

for George C. Marshall Space Flight Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

FOREWORD

This report is one of a series in the field of structural dynamics prepared under contract NAS 8-11486. The series of reports is intended to illustrate methods used to determine parameters required for the design and analysis of flight control systems of space vehicles. Below is a complete list of the reports of the series.

Volume I	Lateral Vibration Modes
Volume II	Determination of Longitudinal Vibration Modes
Volume III	Torsional Vibration Modes
Volume IV	Full Scale Testing for Flight Control Parameters
Volume V	Impedence Testing for Flight Control Parameters
Volume VI	Full Scale Dynamic Testing for Mode Determination
Volume VII	The Dynamics of Liquids in Fixed and Moving Containers
Volume VIII	Atmospheric Disturbances that Affect Flight Control Analysis
Volume IX	The Effect of Liftoff Dynamics on Launch Vehicle Stability and Control
Volume X	Exit Stability
Volume XI	Entry Disturbance and Control
Volume XII	Re-entry Vehicle Landing Ability and Control
Volume XIII	Aerodynamic Model Tests for Control Parameters Determination
Volume XIV	Testing for Booster Propellant Sloshing Parameters
Volume XV	Shell Dynamics with Special Applications to Control Problems

The work was conducted under the direction of Clyde D. Baker and George F. McDonough, Aero Astro Dynamics Laboratory, George C. Marshall Space Flight Center. The General Dynamics Convair Program was conducted under the direction of David R. Lukens.

ACKNOWLEDGMENTS

Portions of this monograph have been taken from internal Convair reports and memoranda not generally available to the reader. In addition, many discussions with various Convair personnel were most helpful to the author. In particular, the discussion in Paragraph 4.2.5.3 is taken from the work of F. I. Backus; his assistance in preparing some of the statistical material in Paragraph 4.3.3.1.2 is greatly appreciated. The derivation of the flexible vehicle equations of motion given in Paragraph 4.3.1.2 was originally due to C. S. Scarborough. Discussions with D. O. Lomen were very helpful in preparing some of the descriptive material on propellant sloshing.

TABLE OF CONTENTS

	<u>Page</u>
List of Illustrations	ix
List of Tables	xi
1 INTRODUCTION	1
2 STATE OF THE ART	3
2.1 Analysis and Synthesis Techniques	3
2.1.1 Frequency Domain Techniques	3
2.1.2 Time Domain Technique	4
2.2 Dynamic Modeling	4
2.2.1 Rigid Body Modeling	5
2.2.2 Propellant Slosh Models	5
2.2.3 Flexible Mode Models	6
2.3 Linear Flight Control System Design	7
3 CRITERIA	9
3.1 Historical Development	9
3.2 Control Authority	10
3.2.1 Thrust Vectoring Capability	10
3.2.2 Autopilot Components	10
3.2.3 Flight Control System Sensors	11
3.2.4 Component Failures	11
3.3 Response	12
3.3.1 Parasitic Modes	12
3.3.2 Limit Cycles	13
3.3.3 Aerodynamic Loads	13
3.3.4 Guidance Commands	15
3.4 Stability	15
3.4.1 Component Tolerances	16
3.4.2 Analytical Uncertainties	16
3.4.3 Response Requirements	16
3.5 Recommendations for Future Development	17

TABLE OF CONTENTS (Continued)

	<u>Page</u>
4 RECOMMENDED PRACTICES	19
4.1 Areas to be Covered	19
4.2 Analysis and Synthesis - Initial Design Phases	19
4.2.1 Rigid Body Flight Control	19
4.2.1.1 Rigid Body Equations of Motion	19
4.2.1.2 Fundamental Design Considerations	22
4.2.1.3 Conventional Autopilot Synthesis for Rigid Body Control	23
4.2.1.3.1 Routh-Hurwitz Criterion	23
4.2.1.3.2 Ultimate State Response	27
4.2.1.3.3 Preliminary Gain Choice	29
4.2.1.4 Direct Synthesis From Standard Forms - Example	33
4.2.2 The Root Locus Method	35
4.2.2.1 Definitions	35
4.2.2.2 Theory	35
4.2.2.3 Construction of Root Loci	37
4.2.2.4 Time and Frequency Response	42
4.2.2.5 Example - Rigid Body Flight Control	45
4.2.3 Frequency Response Methods	46
4.2.3.1 Definitions	47
4.2.3.1.1 Bode Plots	48
4.2.3.1.2 Nichols Charts	48
4.2.3.1.3 Nyquist Diagrams	49
4.2.3.1.4 Mapping	49
4.2.3.2 Nyquist Criterion	49
4.2.3.3 Interpretation of Nyquist Diagrams	53
4.2.3.3.1 Estimation of Roots Having Small Real Parts	53
4.2.3.3.2 Gain and Phase Margins	54
4.2.3.3.3 Complex Gain Margin - Zone of Exclusion	56
4.2.3.3.4 Complex Gain Margin in Root Loci	60
4.2.3.3.5 Extensions of the Nyquist Technique	61
4.2.3.4 Sensitivity to Parameter Variations	62
4.2.3.5 Example - Rigid Vehicle Flight Control	68

TABLE OF CONTENTS (Continued)

	<u>Page</u>
4.2.4 Auxiliary Feedback Loops for Rigid Body Control	74
4.2.4.1 Angle-of-Attack Feedback	76
4.2.4.2 Accelerometer Feedback	76
4.2.4.3 Minimum Drift Gains	78
4.2.4.4 Analysis Using D-Decomposition	78
4.2.4.5 Example - D-Decomposition Analysis of Accelerometer Feedback	81
4.2.5 Stabilization of the Flexible Vehicle Bending Modes	83
4.2.5.1 Equations of Motion	83
4.2.5.2 Initial Choice of Sensor Location and Filter Characteristics	87
4.2.5.3 Approximate Root Location Via Expansion About the Natural Frequency	91
4.2.5.4 Example Case - First-Mode Bending Stability at Liftoff	94
4.2.5.4.1 Selection of Sensor Characteristics	97
4.2.5.4.2 Selection of Engine Position Servo Characteristics	97
4.2.5.4.3 Approximate Root Location	99
4.2.6 Stabilization of Propellant Slosh Modes	103
4.2.6.1 Equations of Motion	103
4.2.6.2 Design Considerations	109
4.2.6.3 Example Case - Slosh Stability Prior to Staging	111
4.2.7 Limit Cycles	113
4.2.7.1 Use of Describing Functions	113
4.2.7.1.1 Thrust Chamber Positioning Servos	114
4.2.7.1.2 Propellant Slosh Damping	117
4.2.7.2 Use of Simulation Techniques	121
4.2.8 General Approach to Preliminary Analysis and Synthesis	121
4.3 Analysis and Synthesis - Detailed Design Phase	123
4.3.1 Development of Equations of Motion	124
4.3.1.1 Basic Principles	124
4.3.1.2 Example - Equations of Motion for Medium Sized, Liquid Propelled Launch Vehicle	126

TABLE OF CONTENTS (Continued)

	<u>Page</u>
4.3.2 Solutions to the Equations of Motion	138
4.3.2.1 Frequency Response	139
4.3.2.1.1 Root Loci	139
4.3.2.1.2 Nyquist Criterion	141
4.3.2.2 Time-Varying Simulation	142
4.3.2.2.1 Analog Computer Simulations	142
4.3.2.2.2 Digital Computer Simulations	143
4.3.2.2.3 Hybrid Computer Simulations	144
4.3.2.2.4 Flight Loads	147
4.3.3 Statistical Influences on System Behavior	147
4.3.3.1 Analytical Techniques	148
4.3.3.1.1 "Worst Case" Approach	148
4.3.3.1.2 Statistical Formulation	148
4.3.3.1.3 Monte Carlo Simulations	151
4.3.3.2 Flight Control System Parameter Tolerances	151
4.3.3.3 Launch Probability and Availability	152
4.4 Correlation of Analytical Predictions and Test Results	153
4.4.1 Production Testing and Preflight Checkout	153
4.4.2 Flight Testing	154
4.5 Summary of Recommended Practice	155
4.6 References	156

LIST OF ILLUSTRATIONS

		<u>Page</u>
1	Rigid Vehicle, Yaw Plane Coordinates and Equations of Motion	20
2	Simplified Flight Control System Block Diagram	24
3	Generalized Block Diagram for Control Systems	36
4	Variations in Root Loci Behavior in the Region of a Saddle Point	40
5	Gain and Phase Margin, Root Locus Interpretation	41
6	Rigid Body Root Locus Under Forward Loop Gain Variation	46
7	Rigid Body Root Locus Under Rate Gain Variation	47
8	The Curve C Used as the Nyquist Criterion for Stability	52
9	Root Location on Nyquist Diagram	54
10	Nyquist Diagram for Conditionally Stable System, Showing Gain and Phase Margin	55
11	Failure of Gain and Phase Margins, as Usually Defined, to Verify System Adequacy	56
12	Complex Gain Margin	58
13	Zone of Exclusion for $K e^{j\phi}$ Having Elliptical Form	59
14	Complex Gain Margin Applied to Root Loci Plots	61
15	Nyquist Diagram of Rigid Vehicle Flight Control System	69
16	Autopilot Block Diagram Showing Physical Placement of Parameters	71
17	Nyquist Diagram of Example System with Zone of Exclusion	74
18	D-Decomposition of K_ψ , K_y Plane for Acceleration Feedback	82
19	Flexible Vehicle, Yaw Plane Coordinates	84
20	Change in Root Locus Departure Angles with Sensor Location, First Bending Mode	89
21	Rigid Vehicle with Propellant Sloshing, Yaw Plane Coordinates and Equations of Motion	104
22	Slosh Zero Location Using Root Locus Technique	107
23	Slosh Pole Shifting Due to Crosscoupling	107

LIST OF ILLUSTRATIONS (Continued)

	<u>Page</u>
24 Stable and Unstable Slosh Pole-Zero Configurations	109
25 Potentially Unstable Limit Cycle, Root Locus	115
26 Potentially Unstable Limit Cycle, Nyquist Diagram	115
27 Stable Limit Cycle	116
28 Propellant Slosh Loci Showing Effect of Increasing Slosh Damping	117
29 Potential Limit Cycle Frequency Versus Slosh Damping	118
30 Position Servo Lag Frequency Versus Engine Deflection Half Amplitude	119
31 Slosh Amplitude Versus Potential Limit Cycle Frequency	119
32 Propellant Slosh Damping Versus Amplitude with Slosh Frequency As a Parameter	120
33 Operating Point for Slosh Limit Cycle	120
34 Coordinate System at Aft End of Launch Vehicle for Modal Deflections, Pitch Plane	129
35 Coordinate System at ℓ^{th} Propellant Tank, Spring-Mass-Couple Analogy	133
36 Block Diagram of Vehicle and Autopilot	140
37 Comparative Solution Times	146

LIST OF TABLES

		<u>Page</u>
1	Rigid Body Parameters for a Medium-Sized Space Launch Vehicle at Maximum Dynamic Pressure	32
2	Flight Control System Parameters - Example Case	32
3	Ranges of Parameter Variation for Example Case - Rigid Vehicle Flight Control	70
4	Evaluation of g_1 , h_1 , and σ_1 - Example Case	73
5	Basic Modal Data for Launch Vehicle at Liftoff - Example Case	95
6	Second-Mode Parameters at Liftoff - Example Case	99
7	"Reduced" Rigid Body and Slosh Parameters for Launch Vehicle Prior to Staging, Example Case	111
8	Slosh Parameter Equivalents	134

1/INTRODUCTION

This monograph presents the techniques in current use for the stability analysis of launch vehicles during their flight through, and upon exit from, the earth's atmosphere. The intent is to indicate the analytical tools and techniques used in the analysis and synthesis of launch vehicle flight control systems during the actively propelled and controlled flight phases. Special considerations such as liftoff, staging, and control during coast flight phases are covered elsewhere in the monograph series.

The object of this volume is to illustrate the applicability of the standard computational techniques to the solution of complex flight control problems associated with large launch vehicles. Current space boosters are characterized, from the flight control standpoint, as large, aerodynamically unstable vehicles whose stability is dominated by parasitic modes of response (elastic vehicle bending and, if liquid propelled, propellant sloshing). Reliance is placed on active flight control by means of thrust vector deflection for stability; fins are used, if at all, as passive elements to decrease the degree of instability in those instances where it proves impractical to adequately stabilize the vehicle with the control system gains allowable by consideration of system parasitics. Also fins may be indicated where abort systems require a reduced rate of attitude divergence in the event of a flight control system failure.

The word "large" in the above context refers to vehicles whose dynamic parameters vary slowly over a wide range. The rates of change of these parameters are far less than those of the dependent variables associated with the short-period modes of response.

The analytical techniques presented are limited to systems that are, or can be approximated by, continuous systems; sampled data systems, digital autopilots, and the like are not considered. There is no coverage of the long-period modes due to guidance system action and guidance commands; rather, the guidance system is treated in a fashion similar to atmospheric disturbances: as a source of system excitation.

2/STATE OF THE ART

The state of the art in linear flight control system design for launch vehicles has reached a high state of development. This has been accomplished using classical techniques for analysis and synthesis. The dynamic modeling of the launch vehicle is one important area where the analyst must rely on engineering judgment, past experience, and ground testing to fill in the details necessary for successful system design.

2.1 ANALYSIS AND SYNTHESIS TECHNIQUES

The analysis and synthesis of launch vehicle flight control systems may be divided into two general areas; that of analysis in the frequency domain, where the properties of the solution to the equations of motion are inferred without actually solving the equations; and in the time domain, where simulations of the launch vehicle are used to obtain solutions to the equations. In addition, there are questions as to the sensitivity of the system to parameter variations and the limit cycle response characteristics resulting from unintentional nonlinearities in the equations of motion.

2.1.1 FREQUENCY DOMAIN TECHNIQUES. The two major techniques in wide use in the analysis of linear systems are the frequency response and the root locus methods. The former has been used in servo system design and analysis for many years and has reached a refined state of development. The basic approach consists of inferring the stability qualities of the closed-loop system from knowledge of its open-loop frequency response. This is accomplished by making use of the Nyquist diagram, which enables the control analyst to determine the relative stability of the system roots from a polar plot of the open-loop frequency response. Various mappings of these plots are in wide use, such as Nichol's charts, Bode plots, and the like; each has certain advantages in reducing the computational load on the engineer.

The root locus method is a somewhat newer technique. Here the location of the roots of the closed-loop system are plotted in the complex frequency domain (s -plane) as a function of various system parameters. The relative stability of the system is determined from the location of the roots relative to the imaginary axis. Since the roots of the system are shown directly, the character of the time response is explicit as to which roots are dominant and to what degree.

Both of these approaches are extremely powerful and well adapted to machine computation. This monograph goes to some length in showing the close relationship between the two: in particular, how one can estimate portions of the Nyquist diagram from the root locus plot and vice versa. Each technique offers its own advantages and disadvantages and should be in the control analyst's "tools of the trade." Exclusive reliance on one or the other can be inconvenient, misleading, or both.

A major extension of the linear theory to the typical nonlinearities encountered in launch vehicle behavior is offered by use of describing functions. This is true in the present application because the nonlinearities are relatively small; the two major sources being the nonlinear damping associated with fluid motion in the launch vehicle propellant tanks and electrohydraulic (or pneumatic) actuators for the thrust vector deflection. This approach can be used either with the root locus or frequency response methods in predicting the possible existence of limit cycles in the vehicle motion.

The stability and response of the launch vehicle are very sensitive to variations in certain system parameters. In order to assure satisfactory behavior under the range of parameter variations to be encountered, one must specify margins in the control system gain and phase characteristics. The margins to be specified are dependent on the sensitivity. Fortunately, many of these quantities can be treated as random variables, thus bringing statistical techniques to bear on the problem. In a great many cases, simple root-sum-squaring of the expected standard deviations of these quantities is sufficient; in others one must use the Monte Carlo approach. Using these methods, the region of possible root locations on the s-plane can be defined as a "zone-of-exclusion" on the Nyquist diagram. From these and studies of the system response properties, suitable gain and phase margins can be defined for each mode in the system response.

2.1.2 TIME DOMAIN TECHNIQUE. Analysis of the launch vehicle in the time domain consists of solving the equations of motion by means of direct simulation of the vehicle dynamics. The latter are complex enough to make inversion of Laplace transforms of the vehicle response impractical except by machine; in addition the parameters are time varying, restricting this technique to short intervals in the vehicle flight history. Nonlinearities can also be incorporated in the simulation, permitting evaluation of the time response of the nonlinear system to command or disturbance inputs.

The detail with which the launch vehicle is simulated depends on the capabilities of the analog, digital, and/or hybrid computers available and the expenditure of money allowable in obtaining a solution. Generally speaking, a comprehensive simulation of the launch vehicle is used to study the behavior in response to guidance commands and aerodynamic disturbances (particularly vehicle loads). In such simulations it is necessary to simulate all major modes of response within the frequency range of the disturbances. In a large, multitank, liquid-propelled vehicle, this would include one or two flexible vehicle modes (perhaps more), the fundamental slosh modes in each tank, and the rigid body mode of response.

2.2 DYNAMIC MODELING

Accurate description of the vehicle dynamics is of paramount importance in the synthesis of a launch vehicle flight control system. The mathematical models used in

various phases of the analytical studies rely to a large extent on past experience and ground testing to augment the purely analytical techniques.

2.2.1 RIGID BODY MODELING. The motion of the launch vehicle at the rigid body control frequency in stability studies is handled on a single-plane, perturbational basis. The symmetry of most launch vehicles is such that there is negligible coupling, from the stability analysis point of view, between the motion about each of the control axes: pitch, yaw, and roll. Such coupling that does exist is normally included only in comprehensive simulations of the launch vehicle. The equations of motion are written in perturbational form; the vehicle flight path changes very slowly in comparison to the frequencies of perturbational motion about the flight path, permitting the decoupling of trajectory-dependent terms from the higher-frequency rigid body motion in stability studies. This coupling is normally included only in detailed simulations of the vehicle motion used for loads studies.

The inertial and aerodynamic characteristics of the launch vehicle are successfully handled using standard techniques. The terms dependent upon the time rate of change of inertial properties are very small and commonly ignored; similarly, quasi-steady aerodynamics are used. The aerodynamic properties are determined on an empirical basis where past experience with configurations of similar geometry are applicable; otherwise model testing in a wind tunnel gives sufficient accuracy.

The class of launch vehicles covered in this monograph obtains the control forces via swiveling of the thrust vector. Most common for liquid-propelled vehicles is gimbaling of one or more thrust chambers. Solid propellant vehicles rely on secondary fluid or gas injection, jetevators, exhaust deflection vanes, and sometimes swiveled nozzles. In all cases, engine inertial properties can be usually ignored in studies of the low-frequency motion of the launch vehicle with the response of the thrust vector deflection angle to autopilot commands represented by a simple lag that may be nonlinear, depending upon the characteristics of the engine positioning servo. The major sources of nonlinearity at these frequencies are gimbal friction and actuator valve dead zone, leakage flow, and mechanical tolerances.

2.2.2 PROPELLANT SLOSH MODELS. In liquid-propelled launch vehicles, the motion of the vehicle in the intermediate frequency range is dominated by propellant slosh forces and the lowest-frequency flexible modes. There are many modes of fluid motion in a rigid walled tank; however, the dominant effect is commonly the first lateral slosh mode. Again, because of vehicle symmetry, there is insignificant coupling between the control axes. Such coupling that does exist is primarily a function of the center-of-gravity lateral offset from the longitudinal axis of symmetry and is included only in detailed simulations. The mathematical techniques for determining the parameters of the lateral mode slosh motion have reached a high state of development. Propellant slosh damping is commonly provided by annular ring baffles in cylindrical tanks, the sizing and spacing of the baffles being determined on both analytical and empirical bases backed up by test results. Mathematically, the forces and

moments produced by the slosh forces are duplicated by spring-mass oscillators or free pendulums, the damping of this motion being a combination of small linear terms and larger, slosh amplitude dependent terms. Flight test data have provided good correlation with this modeling.

2.2.3 FLEXIBLE MODE MODELS. The high-frequency motion of the launch vehicle is due to the free vibrations of structure which, because of the large bandwidth of the flight control system, can be coupled through the autopilot. For stability studies, motion is considered in one plane at a time, the symmetric nature of the vehicle resulting in negligible crosscoupling between the control axes. A lumped parameter mathematical model of vehicle inertial and stiffness properties is used to determine the characteristic modes of vibration*. This approach results in good correlation for the first one or two flexible modes, judging from both ground and flight test results, but is less favorable at the higher modal frequencies. Both the mode shapes and the equations of motion assume that no longitudinal work is done on the structure by the thrust forces, leading to some error as the modal index increases. Fortunately, it is usually possible to gain-stabilize these modes; that is, to decouple them from the flight control system via attenuation within the autopilot filter and control force actuation servo. Thus high precision in the modal parameters for these modes is not ordinarily required.

The equations of motion for the flexible launch vehicle as used in stability studies may be written to include all degrees of freedom, rigid body, propellant slosh, and the flexible modes as well. Such a set of equations is unwieldy and the usual practice is to drop out those degrees of freedom that contribute little to the parasitic mode being investigated. Thus for launch vehicles, where there is as much as an octave frequency separation between the first bending mode and the highest frequency fundamental lateral slosh mode, the latter may be dropped from the equations of motion when the flexible mode stability is being investigated, and vice versa. In time-varying simulations, on the other hand, all degrees of freedom within the frequency range of atmospheric disturbances must be included to get a full picture of the loading conditions on the vehicle structure as it rises through the atmosphere.

When writing the equations of motion to include the effects of propellant slosh and flexible mode bending, the slosh parameters are assumed (for the flexible walled tanks) to be equivalent to those derived for rigid tanks. A spring-mass analogy is used. In most launch vehicles the frequency separation between these two types of parasitic modes is such that the errors involved in this assumption are not too serious, to judge by flight results. Whether this will remain true in vehicles where the frequencies are in closer proximity is another matter.

*R. Gieske, R. Schuett, and D. Lukens, A Monograph on Lateral Vibration Modes, General Dynamics/Convair Report No. GD/C-DDF65-001, 22 February 1965.

In those vehicles that employ gimballed thrust chambers for control, the engine position servo dynamics must be included in flexible mode stability studies; the simple lag approximation no longer suffices because of the increased effect of engine inertial forces at high frequencies. The flexibility of the vehicle structure at the aft end can introduce peculiarities in the response, particularly if there are local, lumped spring-mass systems (turbopumps and the like) that have resonances within the pass band of the positioning servo. The nonlinearities in hydraulic servos at high frequencies are also important and ground testing of the "hot firing" variety is usually required to ensure that the dynamic model of the engine gimbaling system is adequate.

2.3 LINEAR FLIGHT CONTROL SYSTEM DESIGN

As the size, cost, and complexity of space launch vehicles have increased, the sophistication of the flight control system design has necessarily kept pace to meet the increasing demands for control, not only of the rigid body motion to meet guidance objectives, but also to minimize inflight loading on the vehicle structure. The state of the art in flight control system design may be characterized as follows:

The control forces are provided by gimbaling thrust chambers in liquid-propelled rockets and by secondary injection in solids or vehicles using "strap-on" solids. In the former case, the position servo is electrohydraulic, employing a derivative pressure feedback servo valve to linearize the servo response.

The autopilot has continuously programmed gains and usually one or more changes in filtering characteristics during flight to compensate for the changing vehicle parameters. It is computer or programmer controlled and incorporates redundancy features and malfunction detection systems for maximum reliability. The guidance system (which provides the attitude reference) may be inertial or "strap down," depending on mission accuracy requirements. The sensors used include multiple rate gyros in each control channel for either redundancy or flexible mode stabilization or both, and one or more auxiliary sensors (angle of attack meters or accelerometers) for load alleviation during flight. Loads are further minimized by a combination of phase stabilization of the lowest-frequency flexible modes, extensive antislush baffling, and preprogramming of the vehicle boost trajectory to compensate for the winds aloft.

3/CRITERIA

The variety of criteria by which a launch vehicle flight control system may be judged arises from the fundamental requirement of achieving guidance (performance) objectives while maintaining vehicle loads within acceptable levels. The three major areas discussed below; those of control authority, response, and stability criteria; are all aspects of this fundamental criterion. Successful flight control system design takes all these criteria into account together with reliability considerations. The relative importance of each is a function of the range of intended vehicle missions and the vehicle structural load-bearing capabilities.

3.1 HISTORICAL DEVELOPMENT

The criteria used to evaluate the ability of a launch vehicle to satisfy mission requirements have evolved from a wait-and-see philosophy associated with the early relatively unsophisticated vehicles to the highly detailed and proven criteria now in use for the latest and generally largest vehicles. This is in keeping both with the progress in the state of the art and the evolution in mission requirements.

The earliest launch vehicles were intended for use primarily as weapons systems. Here the emphasis was on low unit cost with maximum simplicity in flight control system design. These vehicles were generally not load limited in the interest of having a capability of being launched at any time under any weather conditions. Hence sophistication, in the sense of near optimal flight control system performance, was not a requirement. The design philosophy was that of the simplest possible flight control system that would remain stable throughout flight, limit cycles being allowed if the amplitude was not too great. Response was not a major consideration except in the maximum dynamic pressure region.

The later and more sophisticated launch vehicles intended for space missions are generally load limited to some extent and require a design approach that maximizes launch availability. In addition, the high unit cost per launch requires maximum reliability on a per launch basis; one cannot rely on an extensive R&D flight test program to prove out the vehicle because the cost of such a program is difficult to amortize over the total number of launch vehicles in the program. Consequently, the flight control systems are designed for near optimal performance throughout flight with stability margins maximized within the limits dictated by response requirements. The more sophisticated systems, besides incorporating very conservatively rated components, also include redundancy features and means of detecting and fixing (by automatic substitution) components that fail during flight.

3.2 CONTROL AUTHORITY

Control authority refers to the intended operating range of signal levels within the flight control system. Beyond this range the system can become nonlinear because operation in this range is not a design condition; the vehicle has already violated one or more constraints, usually that of loads.

3.2.1 THRUST VECTORING CAPABILITY. The usual launch vehicle relies on deflection of a portion of total thrust as a means of providing control forces. Control authority, in reference to thrust vector deflection, refers to the required vectoring capability for successful operation. The upper limit on this angle is dictated by the lateral load-bearing capability of the vehicle aft end. The required angle is that necessary to balance the maximum allowable aerodynamic load plus allowance for engine rigging errors, thrust misalignment, lateral center-of-gravity offsets, and dynamic overshoot. Above this point, the allowable load has been exceeded and the vehicle has presumably failed. In short, thrust vector deflection capability should be compatible with the load-sustaining capability (when the vehicle is load limited) or the maximum expected loading (when not load limited).

Physical limitations, such as duct flexibility in gimballed engine installations or the lateral load limit at the aft end of the vehicle structure, may not permit sufficient thrust vectoring capability to be designed into the system. Should this situation arise during preliminary phases of design, the system should be reconfigured to permit sufficient thrust vectoring, by means of an increased control moment arm or whatever other means are available to alleviate the problem. In growth situations where an existing vehicle is being uprated for larger payloads, one must resort to load-relieving schemes that reduce the thrust vectoring requirements. If such a scheme is already being used, one must usually accept a reduction in launch availability. These considerations would indicate the need for considerable foresight in vehicle layout to ensure an adequate thrust vectoring capability throughout the life of the launch vehicle program.

3.2.2 AUTOPILOT COMPONENTS. Generally speaking, autopilot components (servo amplifiers, filters, summing amplifiers, etc.) are designed with a linear range compatible with maximum thrust vector deflection capability. This means that the linear range (defined to be the range of signal levels over which gain, phase, and linearity tolerances are maintained) equals or exceeds the range required to deflect the thrust vector to its limit. One exception to this rule is the autopilot integrator which, when it is separate and distinct from the filtering, may have its output limited to a lower value. The function of the integrator is to permit system operation about an error signal level of zero (neglecting that portion of the error signal required to balance the integrator tendency to drift). The integrator provides an output signal that balances out the "steady state"; that is, the slowly varying errors (thrust misalignment, center-of-gravity offset, rate gyro null offset, and the slowly varying portion of the aerodynamic loads due to the wind profile). In the roll channel, and in a vehicle whose operating environment is outside the atmosphere, the integrator limit need only

be great enough to balance out nonatmospheric effects, while in vehicles that operate in the atmosphere the limit will be compatible with balancing out wind-profile-induced loads as well.

The saturation characteristic of the autopilot integrator is also important. In the saturated (limited) condition, the integrator output voltage ideally should begin decreasing as soon as the input signal causing limiting changes sign; to do otherwise introduces a lag in the system response that is undesirable in the atmospheric environment. Other autopilot components will also exhibit limiting outside their intended operating range and should similarly incorporate no "memory" in the sense of remaining at the limit output value after the input signal begins decreasing.

3.2.3 FLIGHT CONTROL SYSTEM SENSORS. The operating range of the autopilot sensors (rate and position gyros, aerodynamic sensors, etc.) must similarly be compatible with the intended system design. Here the criterion is the expected operating range of the sensed signal. On a launch vehicle that gimbals thrust chambers for control, the engine position transducer will have a linear range equivalent to the full limits of engine travel. The rate gyro, on the other hand, will have a much larger range than expected rigid body rates to allow for the high-frequency signals due to bending deflections. The same is true of accelerometers. Low-frequency sensors such as differential pressure transducers or position sensors will have an operating range more closely compatible with rigid body motion. "Operating range" has the same definition as before: the range of signal levels over which gain, phase, and linearity tolerances are maintained.

3.2.4 COMPONENT FAILURES. A major consideration in system design is that of reliability. Present practice is to specify the operating environment of flight control system components in a very conservative fashion (thermal, shock, and vibration environment) to obtain maximal assurance of normal operation during flight. It is not usually possible to incorporate a high degree of redundancy in booster flight control system design since the allowable range of system parameters is quite narrow. In vehicles whose operating environment is outside the atmosphere, design freedom may be great enough to allow redundancy features. As an example, in a booster incorporating two rate gyros in one control channel the loss of one usually will mean loss of stability. But a deep space vehicle may operate satisfactorily, if poorly, with the rate gain cut in half.

This situation can be alleviated in many cases by designing an automatic monitoring and control system that can sense and act on component failures. This requires detailed knowledge of the possible failure modes and their consequences. When such a system senses a gyro failure it automatically switches out the defective gyro and switches in a standby. The problem becomes one of being able to distinguish between normal and abnormal operation. This might be done by means of a "voting" scheme. One would sense the differences between the output signals of three identical gyros in the pitch channel. When one of these outputs exceeded the other two by some preset

tolerance, the defective gyro would be removed from the loop and the gain on the other two increased to compensate for the loss of the one. Design of such a system requires extensive analysis of the operating characteristics of the hardware and is generally independent of the stability and response considerations.

3.3 RESPONSE

The response properties of the launch vehicle must be compatible with mission and vehicle structural requirements. The most critical situation is the load-limited structure typical of very large boosters. The response must be so tailored as to maintain vehicle loads due to both aerodynamics and parasitic mode response within acceptable levels while still meeting guidance objectives. The fundamental analytical means are frequency response to disturbance inputs and detailed simulation of the launch vehicle flight through the atmosphere. The total load (bending moment) is read out as a function of time and station number and compared with allowables.

In the load-critical situation, response to guidance commands is a secondary consideration; one can correct for booster phase trajectory deviations during the latter stages of flight to a certain extent determined by allowable performance margins. The vehicle loads are of primary concern, and arise from three major sources: aerodynamic environment, parasitic modal response, and limit cycle behavior. The allowable aerodynamic load depends upon the extent to which the latter two can be minimized.

3.3.1 PARASITIC MODES. One contributor to the total vehicle loading is that introduced by vehicle structural flexibility and propellant motion. The vehicle dynamics include sharply tuned (lightly damped) peaks in the frequency response characteristics due to these parasitics. If excited at these frequencies, very large loads can be produced. Fortunately, the sources of excitation can be smoothed (guidance commands) or are random in nature (atmospheric disturbances) such that these resonances are only transiently excited. The problem is then one of damping out the "ringing" of these modes as rapidly as possible.

In the case of slosh parasitics, this can most usually be done by means of antislosh baffles. Not only do these stabilize the parasitic slosh modes, they also reduce the peak slosh forces that can be experienced. However, one can state no formal criteria for the baffling required; there is a tradeoff between the slosh load reduction obtained versus the structural weight of the baffles. The relative importance assigned to each side of this equation is a function of the importance of the slosh loads relative to the overall vehicle loads as well as the criticality of vehicle performance. In the performance-limited vehicle one may trade off launch availability against maximal performance, i. e., minimize vehicle weight and launch only during favorable wind conditions.

Flexible mode parasitics can only be damped with active means within the flight control system. The configuration of autopilot sensors on the vehicle and the phase/gain characteristics of the sensor and autopilot can be so designed as to result in closed-loop flexible mode roots that are relatively well damped. Here the tradeoff is between system complexity with consequent degradation of reliability versus the rapidity with which the modal response dies out. The damping obtainable is rarely more than a few percent, perhaps ten percent of critical at most; but this represents substantial improvement over the inherent damping of the structure which, for the low-frequency modes of concern here, is less than one percent in many cases*.

3.3.2 LIMIT CYCLES. A second source of vehicle loads is that of limit cycle behavior caused by unintentional nonlinearities in the vehicle dynamics or flight control system. There are two major sources of such nonlinearities, propellant slosh damping and the hydraulic (or pneumatic) engine position servo on gimballed engine boosters. The former has been discussed under Paragraph 3.3.1 above, except to relate that a slosh mode, which would be unstable without baffling, will exhibit limit cycles when baffle stabilized; the slosh damping is a function of the slosh amplitude and slosh must be present to obtain the damping. In this case the limit cycle amplitude and its contribution to the vehicle loads are dependent (in part) on the amount of baffling used.

The engine position servo is nonlinear for two reasons: The nonlinear flow/pressure relationship across orifices and the Coulomb friction in the gimbal bearing. The resulting limit cycles will occur at a frequency of a lightly damped mode, typically a slosh mode. If bending and sloshing parasitics are well damped, the limit cycle can occur at the rigid body control mode frequency, particularly if the control mode is lightly damped as it may be after a change in system gain. The cure here is to use compensation within the position servo loop to minimize the nonlinear behavior. The effects of Coulomb friction within the gimbal bearing can largely be eliminated by using a dynamic pressure feedback servo valve. The limit cycles will not be eliminated but can be considerably reduced by this means, particularly if all modes of system response are well damped with considerable phase and (particularly in the case of the conditionally stable rigid body control mode) gain margin.

In both types at limit cycles the criteria of acceptability is, as in the case of parasitic modal response, dependent upon the relative importance of the limit-cycle-induced loads to the total loads. The tradeoff is similar; weight and complexity (with attendant reduction in reliability) are required to minimize the limit cycle loads and may not be justified except in very critical loading situations.

3.3.3 AERODYNAMIC LOADS. The major part of the bending moment experienced by the launch vehicle in its passage through the atmosphere is caused by the aerodynamic load on the vehicle due to angle of attack. Changes in the aerodynamic load

*D. R. Lukens, A Monograph on Full-Scale Dynamic Testing for Mode Determination, General Dynamics/Convair Report No. GD/C-DDF65-002, January 1967.

due to wind gusts further aggravate the loading conditions because of dynamic overshoot in response to these disturbances. In a load-limited vehicle, the flight control system response at rigid body frequencies and below is the major determining factor in the loads experienced. The higher-frequency contributors have already been discussed.

To minimize these loads requires careful design of the rigid body response within the limitations imposed by the need for parasitic modal stability. When using a conventional autopilot (rate and position feedback) a simplified simulation of the rigid body motion can be used to choose gains that minimize the response to gust loads. The slowly varying loads due to the wind profile remain.

Auxiliary feedback loops within the autopilot or guidance system are required to minimize the loads due to the wind profile. These have the property of causing the vehicle to "weather vane" into the relative wind, thus reducing the imposed angle of attack. With proper choice of gains, a "minimum drift" autopilot results. Vehicles with minimum drift autopilots head into the relative wind just enough to balance the wind's effect in blowing the vehicle downstream.

Unfortunately, auxiliary loops also tend to aggravate the vehicle loads at higher frequencies. The vehicle responds faster to a change in the angle of attack, which tends to increase the inertial loads. An optimal choice of the auxiliary loop phase and gain characteristics will minimize the overall loads on the vehicle.

A second means of reducing the aerodynamic loads is via proper preprogramming of the booster flight through the atmosphere. Given knowledge of the wind profile to be flown, the booster pitchover program can be designed to minimize the angle of attack. The aerodynamic loads experienced will depend upon the accuracy with which the wind profile is known and deviations of vehicle performance from nominal.

As in the case of the other contributions to the total vehicle load, the criteria to be applied depend on the particular vehicle flown. If vehicle loading is critical it is likely that both preprogramming of the trajectory and load alleviation features may have to be employed as well as minimizing of the parasitic mode response and limit cycle behavior.

The question of what constitutes reasonable criteria for atmospheric loads; that is, wind criteria; is a subject in itself. Two approaches are used, one an arbitrary selection of "artificial" or "synthetic" wind profiles with superimposed gusts, the second a selection of a number of measured wind profiles likely to produce the greatest loads (tempered with the knowledge of the defects in the measuring process). Unfortunately, both choices suffer from the fact that a worst case or a 95-percent profile, measured or synthetic, is a function of the launch vehicle, although there is strong correlation between the worst wind profiles for any two different configurations. The designer is on firmer ground in using a series of measured profiles together with superimposed

wind gusts to allow for coarseness in measurement. Further, the vehicle simulations flown through these profiles should contain all three control axes. Cases have been found where the profile in each plane is not severe but the combination constitutes the severest condition. This situation is usually associated with a rapid change in the wind direction with the wind speed more or less fixed.

3.3.4 GUIDANCE COMMANDS. The loads acting on the vehicle structure are also dependent upon the form of the guidance (or pitch program) commands. Leaving aside the question of whether or not these commands are tailored for the particular wind profile, the manner in which they are introduced to the vehicle autopilot will influence the loading.

To minimize pitch-program-induced loading, the commands must be "smoothed." For example, if the command signals are in the form of a staircase position command the step size must be relatively small and the frequency large (but not close to a parasitic mode frequency) to minimize the loading. If these conditions are violated one will either excite some parasitic mode to an undesirable high amplitude or induce perturbations in the angle of attack due to overshoot in the rigid body response.

If the guidance commands are in the form of stepwise rate commands torquing a position gyro, the effective position command consists of a series of ramps: a smoother signal than a staircase function. Here the intervals between each step can be considerably longer, several seconds in fact. Even so, the stepwise nature of the pitch program would be detectable in the angle-of-attack history.

The guidance system can also affect the angle of attack because of its drift-minimizing character. Although not covered in this monograph, such systems should be analyzed as part of the overall response problem.

3.4 STABILITY

The stability criteria or stability margins used in the design of the launch vehicle flight control system are composed of two components. The first is the likely range of parameter variations due to analytical uncertainties and component tolerances. The second is the additional margin required for acceptable response of the vehicle. Taking a parasitic mode as an example, it would be insufficient to design for a phase margin of, say, 20 degrees where system variations could use up this margin. An additional 10 or 20 degrees may be required to ensure sufficient damping of the parasitic mode in the vehicle response in the presence of such variations. Defining stability criteria therefore requires knowledge of likely parameter variations and the acceptable region of operating points. Although general stability criteria, such as a 30-degree phase margin and a six- or eight-decibel gain margin, are used they are more or less arbitrary unless qualified by the circumstances pertinent to each mode of response.

3.4.1 COMPONENT TOLERANCES. Depending upon the configuration of the flight control system it is usually possible to maintain variations within the pass band of the control system to within plus or minus one or two decibels in gain and within plus or minus five to twenty degrees in phase, depending upon frequency. The variations from nominal are due to variations in resistor, capacitor, and transistor values, which are random in nature, although not necessarily normally distributed. The non-random operating environment, chiefly thermal, also contributes. Use of high-precision parts, selective assembly, and "tuning" of autopilot elements during manufacture can reduce this range of gain and phase variation somewhat.

Gain and phase variations due to variations of components within their tolerance range can be considered in a statistical fashion, particularly if there is a large number of elements contributing. Under these conditions the range of operating points is almost normally distributed about the operating point, neglecting analytical uncertainty in the description of the vehicle dynamics.

3.4.2 ANALYTICAL UNCERTAINTIES. Analytical uncertainties in the mathematical model can be considered to arise from two sources, the first being due to outright lack of knowledge of the dynamic properties; for example, uncertainty in the modal parameters because of inaccuracies or approximations in the mathematical model. This uncertainty is a fixed, but unknown, quantity and is handled in the analysis via parameter variation over a range considered to cover the maximum possible error that could be made. The location of the first bending mode antinode may not be known with a precision greater than perhaps a quarter of a body diameter, for example.

The second source of uncertainty is interpreted as a statistical quantity: those properties of the vehicle that may vary due to dispersions in the structural properties from vehicle to "identical" vehicle, dispersions in trajectory, performance from launch to launch, etc. The aerodynamic instability parameter might be expected to vary over a five-percent range due to this effect, perhaps another five or ten percent due to model uncertainties. The former would be handled statistically, the latter via parameter variation. Another example of statistical variation is in the slosh modal parameters due to tanking uncertainties and dispersion in propellant burning rates.

3.4.3 RESPONSE REQUIREMENTS. The additional margin required for acceptable system performance in the presence of uncertainties in the system parameters, both fixed and random, is determined from studies of the system response (both transient and steady-state) to environmental disturbances (Subsection 3.3, above). To specify overall gain and phase margins for each type of modal response, control, sloshing, and bending parasitics requires knowledge of the required response. To make an arbitrary statement of gain and phase margin would be meaningless, as cases could always be found that could be overconservative or inadequate for the particular vehicle considered. In final analysis, establishment of such margins requires some measure of engineering judgment, particularly in the area of estimating the fixed analytical uncertainty. The

statistically treated quantities and the required response properties are more amenable to analysis relatively unbiased by the individual engineer's judgment.

3.5 RECOMMENDATIONS FOR FUTURE DEVELOPMENT

Much of the current literature in the field of automatic control is devoted to optimum synthesis procedures. Most of those depend upon an a priori criterion function, which is usually an integral of a quadratic form in various system variables. In the case of a launch vehicle, these variables might be the attitude error, attitude error rate, angle of attack, thrust vector deflection angle, angular rate, lateral drift velocity, etc. By proper selection of the criterion function one can synthesize the flight control system parameters such that the criterion function is a minimum.

The trick is to pick the criterion function. When the flight control system is derived according to the techniques outlined in this monograph, the procedure is basically one of an iterative optimization to certain rule-of-thumb or derived criteria. If a means could be devised by which a criterion function could be selected, the resultant system design would have a firmer analytical basis. The technique would probably come up with a schedule of flight control system gain and phase characteristics impractical to use because of their continuous variation in time. However, the actual gain and phase used could be compared with the ideal as a measure of its optimality with respect to the derived and formalized criterion function. If such a criterion function could be selected for the general case it would be possible to automate the synthesis procedure to a greater extent than is currently practiced, with considerable savings in time and money.

4/RECOMMENDED PRACTICES

This section presents the methods in current use in launch vehicle flight control system analyses of stability and response. It is assumed that the reader is familiar with fundamental control theory; in particular, simultaneous ordinary differential equations, matrix algebra, the Laplace transformation, complex variable theory, and some knowledge of statistics. The derivation is far from complete in the textbook sense and reference is made to standard texts where necessary.

4.1 AREAS TO BE COVERED

The following subsection introduces the major techniques used in the preliminary design phases of launch vehicle development. Simplified equations of motion are given for modes of system response in three frequency ranges: rigid body, propellant sloshing, and flexible vehicle bending. The two major methods of frequency domain analysis, root locus and frequency response, are presented together with a discussion of sensitivity to parameter variations and to phase and gain margins. The subsection concludes with discussions of limit cycle behavior and a general summary of the procedures used in the preliminary design phase. The remaining subsections deal with the design verification phase of analysis, wherein the complexity of the equations of motion necessitate machine solution; a discussion of the correlation between analytical predictions and test results; and a general summary of the recommended practices.

4.2 ANALYSIS AND SYNTHESIS - INITIAL DESIGN PHASES

The stability of space launch vehicles is dominated by the presence of parasitic modes of response in the overall response of the vehicle to guidance commands and atmospheric disturbances. The usual servo system criteria, where the roots of the characteristic equation that determines relative stability are required to have damping ratios between 0.3 and 0.7 of critical, cannot be applied; the parasitic modes are very lightly damped and reliance must be placed on damping ratios of a few percent at most. In many cases these modes are separated enough as a function of frequency to enable separate analysis, at least as a "first cut." This subsection considers rigid body control, elastic vehicle bending parasitics, and propellant sloshing parasitics separately. However, all must be kept in mind when concentrating on one of the three. For example, consideration of bending parasitics will restrict the allowable gains for rigid body control and some acceptable compromise must be found.

4.2.1 RIGID BODY FLIGHT CONTROL

4.2.1.1 Rigid Body Equations of Motion. Figure 1 illustrates the launch vehicle coordinate system and equations of motion in the yaw plane. The equations are for perturbations of the dependent variables about the nominal orientation of the vehicle

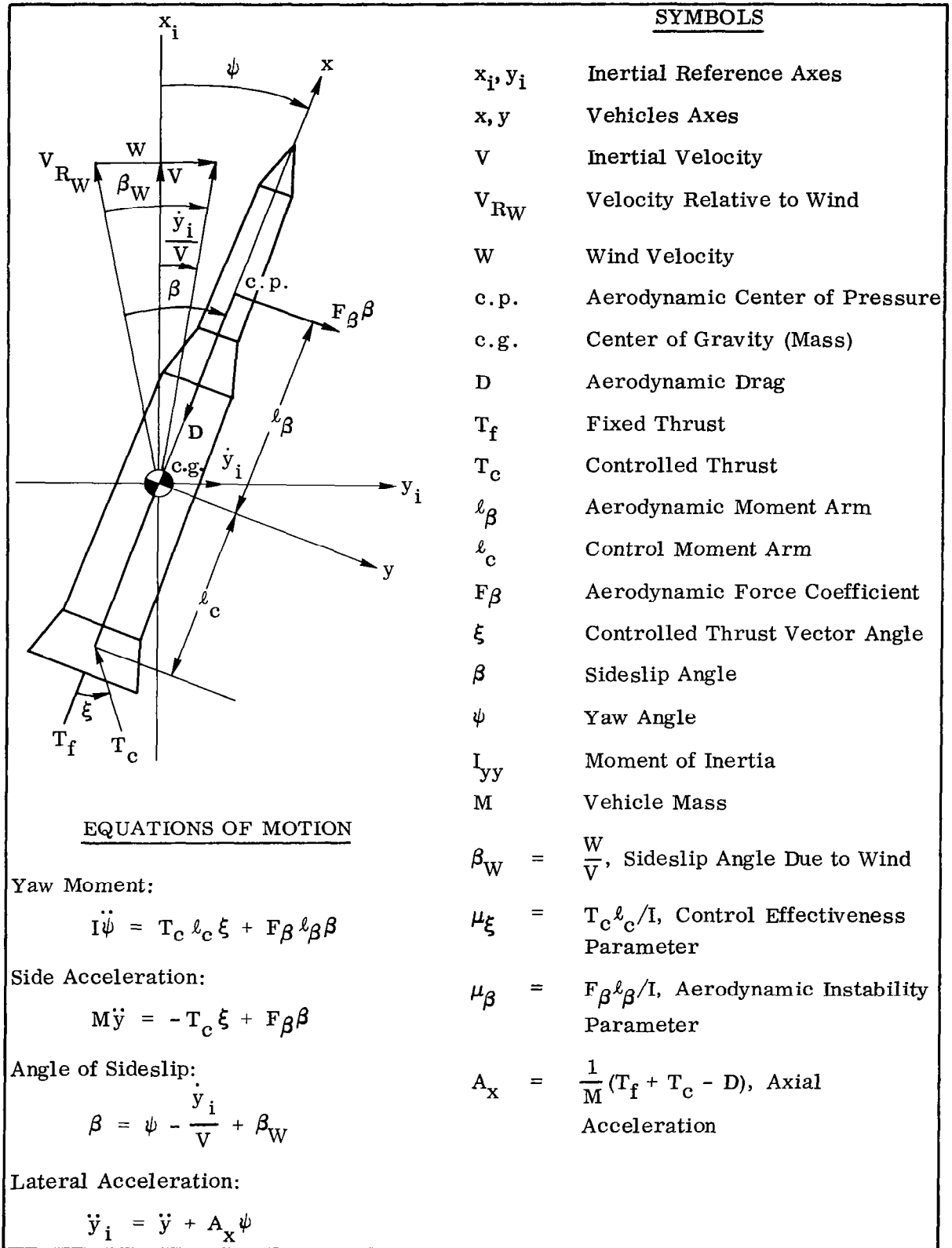


Figure 1. Rigid Vehicle, Yaw Plane Coordinates and Equations of Motion

(the inertial reference frame: x_i, y_i). The booster is aerodynamically unstable with the aerodynamic forces defined along the vehicle axes. Control is obtained by deflecting a portion of the total thrust of the vehicle. The equations of motion can be reduced to the following expression.

$$\begin{bmatrix} s^2 & 0 & -\mu_\beta \\ 0 & 1 & \frac{-F\beta}{M} \\ -(s - \frac{A_x}{V}) & \frac{1}{V} & s \end{bmatrix} \begin{bmatrix} \psi \\ \ddot{y} \\ \beta \end{bmatrix} = \begin{bmatrix} \mu_\xi \\ \frac{T_c}{M} \\ 0 \end{bmatrix} \xi + \begin{bmatrix} 0 \\ 0 \\ s \end{bmatrix} \beta_w \quad (1)$$

This can be resolved into the following transfer function relating vehicle yaw angle to thrust vector deflection by application of Cramer's Rule.

$$\frac{\psi}{\xi}(s) = \frac{\mu_\xi \left[s + \frac{F\beta}{MV} \left(1 + \frac{l_\beta}{l_c} \right) \right]}{s^3 + \frac{F\beta}{MV} s^2 - \mu_\beta s + \frac{\mu_\beta A_x}{V}} \quad (2)$$

At launch, or after exit from the sensible atmosphere, the dynamic pressure is low ($F\beta \rightarrow 0$), leading to the following expression.

$$\frac{\psi}{\xi}(s) = \frac{\mu_\xi}{s^2} \quad (3)$$

At another extreme, when the velocity is very high ($V \rightarrow \infty$) one has

$$\frac{\psi}{\xi}(s) = \frac{\mu_\xi}{s^2 - \mu_\beta} \quad (4)$$

For the large launch vehicles under discussion the latter expression is a good approximation whenever the aerodynamic instability is sufficiently great to necessitate its inclusion. The other terms of Equation 2 lead to a pole-zero dipole in the transfer function, having small influence on the problem (small residue). Physically, the side-slip angle, β , is equal to the yaw angle, ψ , in the absence of β_w .

If differential thrust vector deflection is employed for roll control, the resulting transfer function is similar to Equation 3; thus roll control at all times and pitch/yaw

control when the dynamic pressure is low are a special case of Equation 4. It is assumed, of course, that the launch vehicle is sufficiently symmetric in its aerodynamic properties to result in no appreciable roll moments. Similarly the mass properties are assumed to have sufficient symmetry to result in negligible inertial crosscoupling for purposes of the discussions to follow. Thus control about each vehicle axis may be considered separately.

4.2.1.2 Fundamental Design Considerations. The rigid body control problem is one of stabilizing the unstable pole of Equation 4 in such a way as to lead to acceptable response to guidance commands and, more importantly, to atmospheric disturbances. The latter includes the important consideration of vehicle loads due to aerodynamic and control forces. The loads produced by structural flexibility are also important, but considered later under the heading of parasitic mode stability.

The first consideration is that of the required control authority, i. e., how far the thrust vector must be capable of being deflected to maintain control of vehicle attitude. There are two contributors to the control authority requirements, static (slowly varying) and dynamic (rapidly varying). The thrust vector must be capable of trimming out steady-state (slowly varying) loads due to center-of-gravity offsets, thrust misalignments, and the slowly varying pitch and yaw moments imposed by the wind profiles (horizontal wind speed and direction versus altitude relationships) to be encountered. By "slowly varying" is meant those disturbances whose frequency content is considerably less than the rigid body control frequency. The rapidly varying atmospheric disturbances (predominantly wind gusts, but also severe wind shears) will cause dynamic overshoot in the thrust vector deflection, requiring additional control authority over and above that required for trim conditions.

The launch vehicle is designed for some maximum allowable body bending moment that varies as a function of vehicle station and as a function of time (compressive stresses due to the changing axial acceleration subtract from the bending moment capability; additionally, part of this capability may be dependent on internal pressures that can vary as a function of flight time). The imposed bending moment is composed of steady-state loading due to slowly varying thrust and aerodynamic forces and rapidly changing loads due to atmospheric disturbances and vehicle dynamic response, both rigid body and elastic. The optimum design for a high-performance launch vehicle will have the control authority compatible with the load-sustaining capability, including both static and dynamic effects. This means that the launch vehicle runs out of bending moment capability at or before the thrust vector deflection reaches its maximum.

If sufficient control authority cannot be built into the launch vehicle because of mechanical constraints or lateral load limitations at the aft end of the vehicle, one is faced with three alternatives: the first is to add fins to reduce the aerodynamic moment applied to the vehicle; the second is to incorporate active means within the flight control system to reduce the angles of attack that will be experienced; the third is a combination of the first two.

The second consideration concerns the desired control mode parameters: rigid body frequency and damping ratio. The desired rigid body control response is dominated by the need for minimal (within practicable limits) dynamic overshoot in response to atmospheric disturbances. Response to guidance commands is secondary unless the vehicle is out of the earth's atmosphere, in which case it becomes the controlling requirement. It is axiomatic in control theory that minimal response to disturbances is achieved with high gains and bandwidth in the feedback and compensation signal paths. Unfortunately, this is incompatible with the lightly damped structural modes present on large launch vehicles and an acceptable compromise must be worked out. In the discussion that follows, the selection of gain is based on minimal sensitivity to variations in certain system parameters, but this must be tempered with the knowledge that the gains may have to be even lower for acceptable stability of the elastic modes.

4.2.1.3 Conventional Autopilot Synthesis for Rigid Body Control. The word "conventional" refers to autopilots having proportional plus derivative feedback, and sometimes integral feedback. The proportional term provides the position error required for response to guidance commands; the derivative term is necessary to stabilize the unstable open-loop poles of the transfer function (Equation 4). The integral term is added where it is desired to operate about a zero steady-state position error for reasons of position-sensor or guidance system accuracy. However, this function can be satisfactorily supplied by guidance in some instances. The block diagram of the simplified system, including an actuation lag for the control thrust vector deflection, is shown in Figure 2. The system transfer function is given by

$$\frac{\psi}{\psi_c}(s) = \frac{K_A \mu_\xi (s + K_I)}{\tau_c s^4 + s^3 + (K_A \mu_\xi K_R - \mu_\beta \tau_c) s^2 + [K_A \mu_\xi (1 + K_R K_I) - \mu_\beta] s + K_A \mu_\xi K_I}$$

$$= T(s) \quad (5)$$

where the denominator, when set equal to zero, is the characteristic equation of the system. The roots of this equation, i.e., the roots of the system, determine its stability; the real and complex roots must have negative real parts for the response to an arbitrary input to be bounded in time.

4.2.1.3.1 Routh-Hurwitz Criterion. The Routh-Hurwitz criterion offers a means for determining if any roots of the characteristic equation have positive real parts. Given a characteristic equation of the form

$$a_n s^n + a_{n-1} s^{n-1} + \dots + a_2 s^2 + a_1 s + a_0 = 0 \quad (6)$$

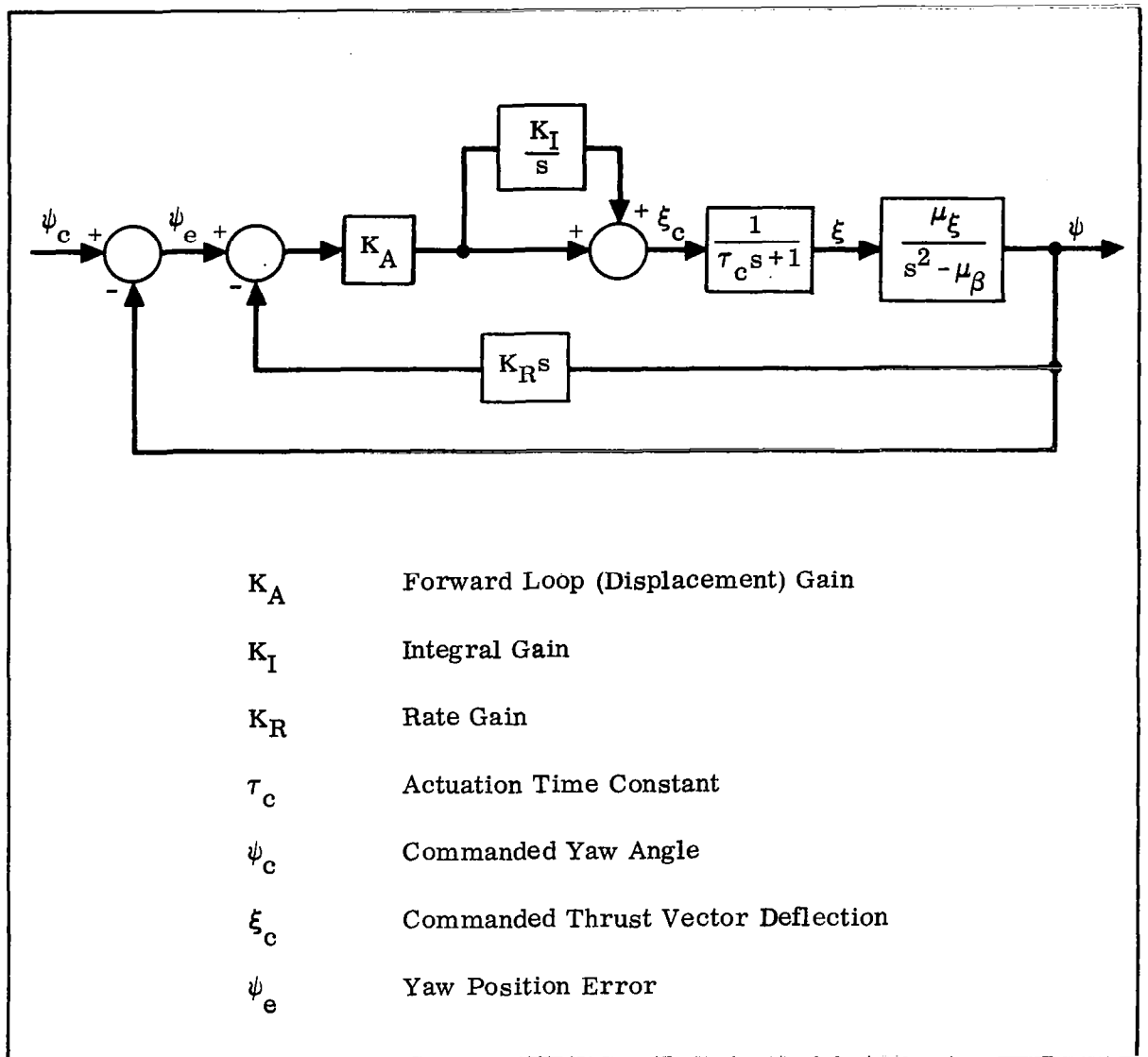


Figure 2. Simplified Flight Control System Block Diagram

an array can be formed, called the Routh array, as follows.

a_n	a_{n-2}	a_{n-4}	\cdot	\cdot	\cdot	\cdot
a_{n-1}	a_{n-3}	a_{n-5}	\cdot	\cdot	\cdot	\cdot
b_1	b_3	b_5	\cdot	\cdot	\cdot	\cdot
c_1	c_3	c_5	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot	\cdot

where

$$\left. \begin{aligned}
 b_1 &= \frac{a_{n-1} a_{n-2} - a_n a_{n-3}}{a_{n-1}} \\
 b_3 &= \frac{a_{n-1} a_{n-4} - a_n a_{n-5}}{a_{n-1}} \\
 b_5 &= \frac{a_{n-1} a_{n-6} - a_n a_{n-7}}{a_{n-1}} \\
 &\vdots \\
 c_1 &= \frac{b_1 a_{n-3} - b_3 a_{n-1}}{b_1} \\
 c_3 &= \frac{b_1 a_{n-5} - b_5 a_{n-1}}{b_1} \\
 d_1 &= \frac{c_1 b_3 - b_1 c_3}{c_1} \\
 d_3 &= \frac{c_1 b_5 - b_1 c_5}{c_1} \\
 &\vdots
 \end{aligned} \right\} \quad (7)$$

The array when complete will consist of $n + 1$ rows, where the number of non-vanishing terms in each row decreases by one following every second row, beginning with the first or second row, as n , the degree of the equation, is even or odd. If the first form in any row vanishes, it is replaced by an arbitrarily small constant, ϵ , of the same sign as the first term in the preceding row (to prevent the terms in the following row from becoming infinite).

The criterion is stated as follows: if the array is complete; if none of the first $n + 1$ rows contain only zeros and none of the elements of the first column (the initial terms in each row) vanish; then the system is stable if all elements of the first column are of

the same sign. If there are changes in sign, then the number of roots having positive real parts is given by the number of changes in sign. If one of the elements vanishes, it indicates a root or roots having a zero real part, which, by the definition of stability given earlier, indicates an unstable system. It should be noted that any characteristic equation having coefficients of differing sign will always result in a Routh array having sign changes in the first column; hence one knows immediately that the system is unstable.

To apply this criterion to the system of Equation 5 one forms the array as follows.

$$\begin{array}{ccc}
 \tau_c & (K_A \mu_\xi K_R - \mu_\beta \tau_c) & K_A \mu_\xi K_I \\
 1 & [K_A \mu_\xi (1 + K_R K_I) - \mu_\beta] & \\
 b_1 & K_A \mu_\xi K_I & \\
 c_1 & & \\
 K_A \mu_\xi K_I & &
 \end{array}$$

where

$$\left. \begin{aligned}
 b_1 &= \tau_c [K_A \mu_\xi (1 + K_R K_I) - \mu_\beta] - (K_A \mu_\xi K_R - \mu_\beta \tau_c) \\
 c_1 &= (K_A \mu_\xi K_R - \mu_\beta \tau_c) - \frac{K_A \mu_\xi K_I}{\tau_c [K_A \mu_\xi (1 + K_R K_I) - \mu_\beta] - (K_A \mu_\xi K_R - \mu_\beta \tau_c)}
 \end{aligned} \right\} \quad (8)$$

Note that the last term in every second row repeats itself, a useful check when applying the criterion in numerical form. Since the first, second, and fifth elements are positive, the Routh-Hurwitz criterion then requires that b_1 and c_1 be positive. This results in two conditions derivable from the above expressions:

$$\left. \begin{aligned}
 \frac{K_R}{1 + K_R K_I} &> \tau_c \\
 K_A \mu_\xi (1 + K_R K_I) - \mu_\beta &> \frac{\frac{K_I}{1 + K_R K_I}}{\frac{K_R}{1 + K_R K_I} - \tau_c}
 \end{aligned} \right\} \quad (9)$$

The first of these criteria states that the ratio of end-to-end rate gain ($\dot{\psi}$ to $\dot{\xi}_c$) to end-to-end position gain (ψ to ξ_c) must exceed the actuation lag. The second says that the control moment per unit attitude change must exceed the upsetting moment by a quantity equivalent to the ratio between the end-to-end integral gain and the amount by which the first criterion is satisfied. If there is no integrator present, these criteria reduce to the simpler expressions:

$$K_A \mu_\xi > \mu_\beta \quad \text{and} \quad K_R > \tau_c \quad (10)$$

The interpretation is clear: system lead must overcome system lag and the control moment produced by an attitude error must exceed the aerodynamic upsetting moment causing that error.

The major conclusion to be drawn from this discussion is that application of the Routh-Hurwitz criterion becomes quite cumbersome for application to systems of even moderate order, especially if an explicit interpretation in terms of system gains is to be obtained. Secondly, it gives no information as to relative stability of the system, merely that the roots of the system have negative real parts. The technique can be extended in a limited fashion by substituting the quantity, $s + \sigma$, for the Laplace transform variable, s , in the characteristic equation. Application of the criteria then reveals whether or not the roots of the equation have real parts more negative than σ .

4.2.1.3.2 Ultimate State Response. One consideration in the design of the flight control system is the steady-state error in response to guidance commands. In the system of Figure 2 the integrator in the error channel guarantees zero steady-state error in response to a constant $\psi_c(t)$. This can be determined by inspection from Equation 5 where the lowest order (constant) terms in numerator and denominator are equal. This follows from the final value theorem, which states for stable systems that

$$\lim_{t \rightarrow \infty} [\psi(t)] = \lim_{s \rightarrow 0} [s \psi(s)] \quad (11)$$

The input is taken to be a unit step, $\psi_c(s) = 1/s$, and for the system of Equation 5 the final value is the ratio of the constant terms in numerator and denominator; in this case, unity.

Evaluation of the response to a constant commanded turning rate, $\psi_c(s) = 1/s^2$, is somewhat more complicated. The output is broken into a partial fraction expansion:

$$\begin{aligned} \psi(s) &= \frac{T(s)}{s^2} \\ &= \frac{A_1}{s^2} + \frac{A_2}{s} + \frac{F_1(s)}{F_2(s)} \end{aligned} \quad (12)$$

where

$$\left. \begin{aligned} A_1 &= T(s) \Big|_{s=0} \\ A_2 &= \frac{d}{ds} T(s) \Big|_{s=0} \end{aligned} \right\} \quad (13)$$

and the remaining terms represent that portion of the response which tends to zero as $t \rightarrow \infty$, providing, of course, that the system is stable. If

$$T(s) = \frac{a_m s^m + a_{m-1} s^{m-1} + \dots + a_2 s^2 + a_1 s + a_0}{b_n s^n + b_{n-1} s^{n-1} + \dots + b_2 s^2 + b_1 s + b_0} \quad (14)$$

where $n \geq m$, then

$$\left. \begin{aligned} A_1 &= \frac{a_0}{b_0} \\ A_2 &= \frac{a_1 b_0 - a_0 b_1}{b_0^2} \end{aligned} \right\} \quad (15)$$

In the case of Equation 5, one has

$$\left. \begin{aligned} A_1 &= 1 \\ A_2 &= - \left(K_R - \frac{\mu_\beta}{K_A \mu_\xi K_I} \right) \end{aligned} \right\} \quad (16)$$

This states that the system will follow a ramp input with zero velocity error, but lagging in time by the quantity, $K_R - \frac{\mu_\beta}{K_A \mu_\xi K_I}$. This term is a frequent source of discrepancy between trajectory simulations employing a point-mass vehicle and instantaneous autopilot and one containing a model of the low-frequency vehicle dynamics.

The expressions of Equations 15 also indicate the possibility of following a ramp input with zero steady-state error when the launch vehicle has left the atmosphere ($\mu_\beta \rightarrow 0$). If the feedback rate signal is introduced downstream of the integrator, the

first-order term in the denominator becomes equal to the first-order term in the numerator and A_2 vanishes. If this scheme is employed during the ascent through the atmosphere, the vehicle position will tend to lead the commanded position ($A_2 > 0$); in fact, it will probably lead anyway since K_I is normally a relatively small quantity.

4.2.1.3.3 Preliminary Gain Choice. The major consideration in flight control system design is the minimization of vehicle loads caused by atmospheric disturbances. The conventional autopilot under discussion here will not respond to a wind gust until an attitude change has occurred. This change will always be in a direction to increase the angle of attack for an aerodynamically unstable vehicle. In illustration of this, consider a simplified flight control system with rate and position gain only, the actuation lag and integrator being neglected. The transfer function relating β to β_W becomes

$$\frac{\beta}{\beta_W}(s) = \frac{s^2 + K_A \mu_\xi K_R s + K_A \mu_\xi}{s^2 + K_A \mu_\xi K_R s + (K_A \mu_\xi - \mu_\beta)} \quad (17)$$

where the denominator, when set equal to zero, is the characteristic equation for the system. The time response to a unit step input (unrealistic for gust inputs, but convenient) is given by

$$\beta(t) = 1 + \frac{\mu_\beta}{\omega_{RB}^2} \left\{ 1 - \frac{e^{-\zeta_{RB} \omega_{RB} t}}{\sqrt{1 - \zeta_{RB}^2}} \sin \left[\omega_{RB} \sqrt{1 - \zeta_{RB}^2} t + \tan^{-1} \frac{\sqrt{1 - \zeta_{RB}^2}}{\zeta_{RB}} \right] \right\} \quad (18)$$

where $\omega_{RB} = \sqrt{K_A \mu_\xi - \mu_\beta}$, the undamped closed-loop natural frequency, and $\zeta_{RB} = \frac{K_A \mu_\xi K_R}{2 \omega_{RB}}$, the system damping ratio. The maximum value of β occurs at

$t = \frac{\pi}{\omega_{RB} \sqrt{1 - \zeta_{RB}^2}}$ and is given by

$$\beta_{\max} = 1 + \frac{\mu_\beta}{\omega_{RB}^2} \left[1 + e^{-\frac{\zeta_{RB} \pi}{\sqrt{1 - \zeta_{RB}^2}}} \right] \quad (19)$$

It is clear that the overshoot, hence the vehicle aerodynamic loading, is reduced as the gain and bandwidth of the autopilot are increased. However, neither can be increased indefinitely. Apart from stability considerations of the parasitic modes, the vehicle loads will begin increasing again as the gain is raised due to increasing vehicle

inertial loads as the response speed becomes greater. A convenient means to demonstrate this is to set up a simplified simulation on an analog computer capable of reading out body bending moment at critical stations along the vehicle. The test input can be a more realistic input of the form:

$$\beta_W(t) = \begin{cases} 1 - \cos \omega_W t, & 0 \leq t \leq \frac{2\pi}{\omega_W} \\ 0 & t \geq \frac{2\pi}{\omega_W} \end{cases} \quad (20)$$

At some intermediate point the gains will be such as to minimize the load for a particular gust frequency, ω_W . The problem then becomes one of choosing what constitutes realistic gust frequencies. This gets into the question of wind statistics which is beyond the limited scope of the present discussion. The criteria for preliminary choice of system gains must be based on more limited criteria that are easier to apply.

One such consideration is that of sensitivity to system parameter variations. Both autopilot gains and the degree of aerodynamic instability are subject to variations, the former due to tolerances within the autopilot and the latter due to variations in launch vehicle performance and trajectory.

A convenient rule of thumb is to choose the displacement gain twice the ratio of μ_β/μ_ξ . This is based on minimizing the expected variations in the rigid body damping, ζ_{RB} , as K_A (and/or μ_ξ) varies. One has

$$\begin{aligned} \frac{\partial \zeta_{RB}}{\partial K_A} &= \frac{\partial}{\partial K_A} \left[\frac{K_A \mu_\xi K_R}{2\sqrt{K_A \mu_\xi - \mu_\beta}} \right] \\ &= \frac{K_R \mu_\xi (K_A \mu_\xi - 2\mu_\beta)}{4 (K_A \mu_\xi - \mu_\beta)^{3/2}} \end{aligned} \quad (21)$$

which is zero for $K_A \mu_\xi = 2\mu_\beta$. This also sets the natural frequency of the closed-loop system at $\omega_{RB} = \sqrt{\mu_\beta}$ for the simplified approximation here.

The choice of rate gain is equivalent to the choice of the damping ratio of the system. Given the above choice for displacement gain, the major parameters affecting the damping ratio are the rate gain and aerodynamic instability parameter, μ_β . Applying the same technique as above, one has

$$\left. \begin{aligned} \frac{\partial \zeta_{RB}}{\partial K_R} &= \frac{K_A \mu_\xi}{2 K_A \mu_\xi} \\ \frac{\partial \zeta_{RB}}{\partial \mu_\beta} &= \frac{K_A \mu_\xi K_R}{4 (K_A \mu_\xi - \mu_\beta)} \end{aligned} \right\} \quad (22)$$

The expected change in ζ_{RB} is therefore

$$\Delta \zeta = \zeta_{nom} \left(\frac{\Delta K_R}{K_R} \right) + \frac{\mu_\beta}{2 \omega_{RB}^2} \zeta_{nom} \left(\frac{\Delta \mu_\beta}{\mu_\beta} \right) \quad (23)$$

If the minimum value of ζ_{nom} is selected as ζ_{min} , one obtains

$$\zeta_{nom} = \zeta_{min} + \Delta \zeta$$

The choice for ζ_{min} is dictated by the degree of stability required with due considerations for the system lags that have been neglected up to this point.

The actuator lag τ_c cannot be reduced indefinitely because of practical power and structural limitations. While a small time constant is desirable in reducing the amount of lead necessary to stabilize the system, the practical problems associated with gimbaling a massive thrust chamber (should this means be used for thrust vector deflection) would certainly limit the reduction possible in τ_c . On the other hand, the philosophy wherein τ_c is made as small as practically possible has merit in that it allows considerable flexibility in the design of the system. If additional lag is necessary for stabilization of vehicle parasitic modes of response, it can most easily be added electrically. Overcoming the limitations of a sluggish actuation system is considerably more difficult from the practical standpoint. A rule-of-thumb value for $1/\tau_c$ is ten times the rigid body natural frequency:

$$\frac{1}{\tau_c} = 10 (K_A \mu_\xi - \mu_\beta) = 10 \omega_{RB} \quad (25)$$

This assures minimal influence on the rigid body control mode.

Use of an autopilot integrator is indicated where it is desired to operate about a zero steady-state error signal for reasons of guidance system or position reference limitations. A small value is desirable for minimal effect on the rigid body control mode. Further, only a small value is necessary since the integrator only nulls out the slowly varying parameters affecting trim conditions (thrust misalignment, center-of-gravity shifts, and the like). A reasonable choice is

$$K_I \leq \frac{1}{10} \sqrt{K_A \mu_\xi - \mu_\beta} \quad (26)$$

It would be helpful at this point to take a typical example of a launch vehicle, work out a choice of gains, and obtain a measure of the approximations made in their choice. Table 1 lists the pertinent parameters. For this case, the rate gain is based on $\zeta_{\min} = 0.4$ for 20-percent variations in K_R and μ_β , resulting in the system gains of Table 2. The degree of approximation made earlier, that the velocity is assumed infinite, may be seen by substituting the values in Table 1 into the transfer function of Equation 2. The aerodynamic roots are shifted to the left by about one percent, and a pole-zero dipole appears at the origin. The approximations made in the choice of gains (neglecting integrator and actuation lags) may be verified by substituting the values given into the characteristic equation (denominator of Equation 5 set equal to zero). We obtain

$$s^4 + 18.52 s^3 + 35.80 s^2 + 70.7 s + 23.4 = 0 \quad (27)$$

which yields the factors:

$$(s + 0.394)(s + 16.616)(s^2 + 1.510 s + 3.575) = 0 \quad (28)$$

Table 1. Rigid Body Parameters for a Medium-Sized Space Launch Vehicle at Maximum Dynamic Pressure

$F_\beta = 298,000 \text{ lb/rad}$
$l_\beta = 36.3 \text{ ft}$
$l_c = 34.1 \text{ ft}$
$M = 5720 \text{ slugs}$
$\mu_\beta = 3.407 \text{ sec}^{-2}$
$\mu_\xi = 3.901 \text{ sec}^{-2}$
$T_c = 362,000 \text{ lb}$
$V = 1709 \text{ ft/sec}$
$A_x = 69.3 \text{ ft/sec}^2$

Table 2. Flight Control System Parameters - Example Case

$$K_A = 2 \frac{\mu_\beta}{\mu_\xi} = 1.75$$

$$\tau_c = \frac{1}{10 \mu_\beta} = 0.054 \text{ sec}$$

$$K_I = \frac{1}{10} \mu_\beta = 0.185 \text{ sec}^{-1}$$

$$K_R = \frac{\zeta_{\text{nom}}}{\mu_\beta} = 0.310 \text{ sec}^*$$

$$(\zeta_{\text{nom}} = 0.571)$$

*After one iteration on the full characteristic equation (including integrator and actuator lags), an estimate of $K_R = 0.45 \text{ sec}$ is achieved.

Thus a natural frequency, $\omega_{RB} = 1.89$ rad/sec, and damping, $\zeta_{RB} = 0.40$, result. This compares with the intended values of 1.85 rad/sec and 0.571, respectively. In this case, the rate gain was not chosen high enough to give the desired nominal ratio of 0.571. Making use of the first term on the right-hand side of Equation 23, it is estimated that an increase of K_R to approximately 0.45 sec would be required to bring ζ_{RB} to its intended nominal value. At this point, the preliminary gain choices for rigid body control at maximum dynamic pressure are complete.

The first approximation gains chosen above should form a satisfactory basis for more elaborate analyses. Before proceeding to synthesis using frequency response and root locus techniques, one major consideration should be stressed. The properties of a launch vehicle as to propulsive, aerodynamic, and inertial parameters vary over a wide range through the boost phase; the gain selection here will probably have to be modified for other flight times, especially when considering the parasitic modes of response.

4.2.1.4 Direct Synthesis From Standard Forms - Example. After the launch vehicle has left the sensible atmosphere, the design of the vehicle flight control system is dominated by the requirements for response to guidance commands. The upper stages of a launch vehicle thus have considerably less stringent requirements on the flight control system, the major system disturbance being that of engine ignition. The flight control system gains can be quite low with very rapid roll-off in the open-loop response. Further, the system can be designed for optimal response with respect to formalized criteria.

One such criterion is the ITAE, or integral of the time-multiplied absolute error, given by

$$C_{ITAE} = \int_0^{\infty} t |\psi - \psi_c| dt \quad (29)$$

Application of this criterion results in reasonably well-damped systems having overshoots that are not too severe. This is in contrast to the usual integral square error criterion.

A second desirable requirement is that the system respond to the commanded rates with zero velocity error. This is accomplished by introducing the rate signal downstream of the autopilot integrator, resulting in the closed-loop transfer function:

$$\frac{\psi}{\psi_c}(s) = \frac{K_A \mu_{\xi} (s + K_I)}{s^3 + K_A \mu_{\xi} K_R s^2 + K_A \mu_{\xi} s + K_A \mu_{\xi} K_I} \quad (30)$$

The optimum (in the ITAE sense) configuration for the denominator polynomial is given by

$$s^3 + 1.75 \omega_0 s^2 + 3.25 \omega_0^2 s + \omega_0^3 \quad (31)$$

This factors into

$$(s + 0.220 \omega_{RB})(s^2 + 0.836 \omega_{RB} s + \omega_{RB}^2) \quad (32)$$

where the parameter is now the undamped natural frequency for the rigid vehicle control mode and which is chosen based on response requirements to guidance commands, presumably quite low, perhaps 0.5 rad/sec. Equating coefficients, one finds

$$\left. \begin{aligned} K_A \mu_\xi &= 1.185 \omega_{RB}^2 \\ K_R &= 0.893 \omega_{RB}^{-1} \\ K_I &= 0.1855 \omega_{RB} \end{aligned} \right\} \quad (33)$$

If the actuation lag is chosen high enough to have negligible influence on the three dominant roots, say $5 \leq (\tau_c \omega_{RB})^{-1} \leq 10$, the system has been synthesized. One might question the utility of the fourth-order standard form (zero velocity error) for the ITAE criteria, whose characteristic equation is

$$s^4 + 2.41 \omega_0 s^3 + 4.93 \omega_0^2 s^2 + 5.14 \omega_0^3 s + \omega_0^4 = 0 \quad (34)$$

In terms of the rigid body natural frequency,

$$(s + 0.691 \omega_{RB})(s + 0.1366 \omega_{RB})(s^2 + 0.508 \omega_{RB} s + \omega_{RB}^2) \quad (35)$$

Unfortunately, the rigid body damping is low enough to make the maintenance of adequate stability margins questionable, even though the actuation lag can be quite low. In practice the best approach in this instance would be to set up a simplified system on the analog computer so as to read out Equation 29 as a function of time. Then, constraining the actuation lag and rigid body control frequency to fixed values, one would vary the dominant root damping ratio and the fourth unconstrained root to arrive at a minimum C_{ITAE} .

The foregoing is intended as an example of the application of standard forms for synthesis. At best they can be used as a guide in those situations that are ideal. The

design can start with those parameters tabulated in literature (Reference 1) but it must be investigated for compliance to other system constraints and changing behavior with flight time.

4.2.2 THE ROOT LOCUS METHOD. One of the most powerful tools applicable to the analysis and synthesis of launch vehicle flight control systems is the root locus technique. Its chief virtue is the graphical display of the motion of system roots as functions of various system parameters. This enables the control engineer to evaluate the effects of these changes directly (the root locations can be translated immediately into an estimate of the system time response) or if desired, its frequency response. The following discussion is an exposition on the root locus method where the theoretical background is considerably abbreviated. It can be found in varying states of clarity and completeness in almost any textbook on linear control theory; for example, References 2 and 3.

4.2.2.1 Definitions. The roots of the characteristic equation are a set of complex numbers that depend upon the coefficients of the equation. These roots are continuous functions of the coefficients, hence of the system parameters. If the roots of the characteristic equation are plotted on the complex s -plane, they will trace out continuous curves as the several system parameters are varied. The path followed for a particular root is a root locus. For a given set of parameter values there is a single set of roots but through these roots a different set of loci will pass for each parameter variation. The question is then one of determining these loci.

4.2.2.2 Theory. The characteristic equation of a linear control system may be written in the form:

$$1 + GH(s) = 1 + K \frac{\prod_{i=1}^m (s - z_i)}{\prod_{j=1}^n (s - p_j)} = 0 \quad (36)$$

where the z_i and p_j are respectively the zeros and poles of the system, both real and complex, and K is the "high-frequency open-loop gain" — proportional to the low-frequency system gain. The function $GH(s)$ is the open-loop transfer function for the system (see Figure 3). Suppose the complex number, s_0 , is a root of Equation 36. Then, since gain K is a real number, the characteristic equation is equivalent to two relations:

$$K = \frac{\prod_{j=1}^n |s_0 - p_j|}{\prod_{i=1}^m |s_0 - z_i|} \quad (37)$$

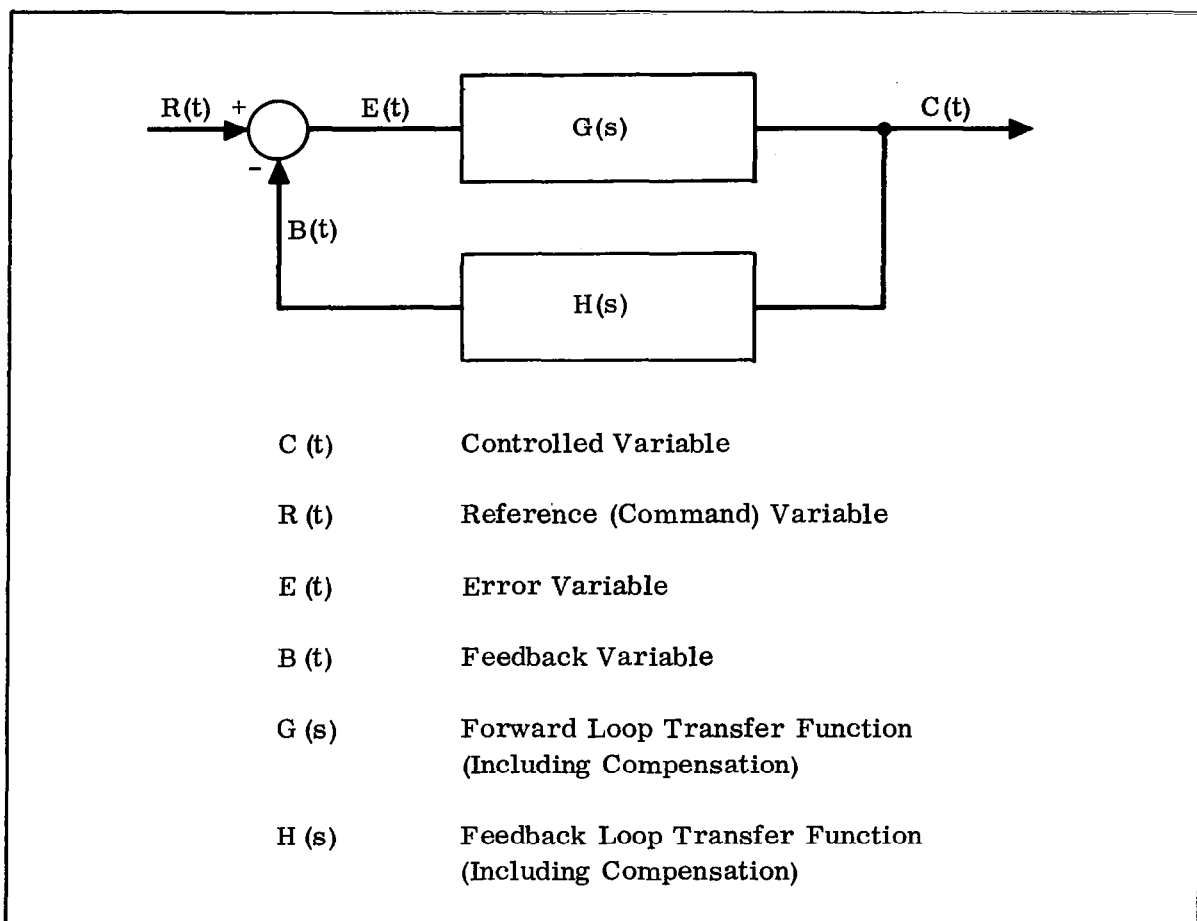


Figure 3. Generalized Block Diagram for Control Systems

$$\text{Arg} \left[\frac{\prod_{i=1}^m (s_0 - z_i)}{\prod_{j=1}^n (s_0 - p_j)} \right] = \sum_{i=1}^m \text{Arg} (s_0 - z_i) - \sum_{j=1}^n \text{Arg} (s_0 - p_j)$$

$$= (2r + 1)\pi \quad (38)$$

where r is any integer. If a point can be found in the s -plane satisfying Equation 38, there will be only one value of K for which Equation 37 is satisfied. The root locus for s_0 under variation of the open-loop gain is the locus of points satisfying Equation 38 with K a parameter.

4.2.2.3 Construction of Root Loci. The procedure for plotting root loci is as follows:

1. Plot poles p_j and zeros z_i of the open-loop transfer function $GH(s)$. The convention normally adopted is that of using "o" for zeros, "x" for poles. The real and imaginary axis of the s-plane plot must have the same linear scale.
2. Locate the point on the real axis given by

$$s = \frac{\sum_{j=1}^n p_j - \sum_{i=1}^m z_i}{n - m} \quad (39)$$

which is sometimes referred to as the "centroid" of the pole-zero configuration.

3. Draw the asymptotes that the root loci will follow as $K \rightarrow \infty$, given by radial lines intersecting at the centroid and making angles given by (measured counterclockwise from the 3 o'clock position):

$$\theta = \begin{cases} (2r + 1)\pi / (n - m) & \text{(negative feedback)} \\ 2r\pi / (n - m) & \text{(positive feedback)} \end{cases} \quad (r = 1, 2, \dots) \quad (40)$$

If the number of poles equals the number of zeros (i.e., $n = m$) there are no asymptotes and this step is omitted.

4. Determine those parts of the real axis containing a portion of the root loci according to the following rule: loci lie on the real axis to the left of an odd number of poles and zeros for negative feedback and to the right for positive feedback.
5. Determine the departure angles from the poles and approach angles to the zeros as follows:
 - a. For single poles or zeros on the real axis, step 4 satisfies this requirement.
 - b. For double poles or zeros on the real axis, step 4 satisfies or the loci depart perpendicularly from the real axis.
 - c. For triple poles (generally found, if at all, at the origin of the s-plane for flight control systems), step 4 defines one of the loci, and the other two make angles of 120 degrees with respect to it.
 - d. For complex poles the loci depart at an angle θ given by

$$\theta = (2r + 1)\pi - \sum_{j=1}^n \text{Arg}(p_k - p_j) + \sum_{i=1}^m \text{Arg}(p_k - z_i) \quad (41)$$

where the angle is measured counterclockwise from the 3 o'clock position.

e. For complex zeros, the loci approach at an angle given by

$$\theta = (2r + 1)\pi + \sum_{j=1}^n \text{Arg}(z_k - p_j) - \sum_{i=1}^m \text{Arg}(z_k - z_i) \quad (42)$$

where the angle is measured counterclockwise from the 3 o'clock position.

In these relations, $\text{Arg}(p_k - p_j)$ represents the angle made by a vector drawn from p_j to p_k as measured from the three o'clock position at p_j and similarly for the other angles.

6. Determine points σ_B where roots meet on the real axis and break away perpendicularly according to a trial and error solution of

$$\sum_{i=1}^m \text{Re} \left[\frac{1}{\sigma_B - z_i} \right] = \sum_{j=1}^n \text{Re} \left[\frac{1}{\sigma_B - p_j} \right] \quad (43)$$

where the operator, Re , signifies "the real part of."

7. Sketch in the root loci, bearing in mind that the root loci are similar to stream lines in a potential flow problem where the poles represent sources and the zeros represent sinks. Alternatively, the root locus between a pole and a zero may be thought of as the motion of an electron emitted from a line cathode (the pole) as it travels to the line anode (the zero). In short, a locus bends away from poles (and other loci) and toward zeros.
8. Complete the plot using the sketched loci as guides. This amounts to picking a test point, s_0 , near the sketched loci and measuring the angles of Equation 38 to see if the equation is satisfied; a trial and error process. This procedure will be greatly facilitated by using a Spirule* that accumulates a running total of the angles.

* A Spirule can be obtained from: The Spirule Company
9728 El Venado
Whittier, California

together with explicit and detailed instructions on its use. Although the procedure outlined here reads as though it were quite laborious, it is in fact quite rapid; with a little practice the plots can be obtained quickly.

9. Determine the gain variation along the loci. This is done by measuring (in units of s ; if the plot has 0.5 inch equaling 1 rad/sec, 3.75 inches represents 7.5 radians) the distances between a point s_0 on the locus where the gain is to be determined and all the poles p_j and zeros z_i ; then determine K according to Equation 37. This can also be done using a Spirule. In the near vicinity of a pole the gain will be found to vary in direct proportion to the distance to that pole, with the other distances in Equation 37 changing only slightly. In the vicinity of a zero the gain will vary inversely proportional to the distance to the zero. The words "near vicinity" mean that the test point is near the pole (or zero) relative to its distance from any other pole or zero.

This procedure will define the motion of the system roots as a function of K . A few observations are in order. On the s -plane there may exist "saddle points," points where the derivative of the complex gain function given by

$$\frac{d}{ds} K(s) = - \frac{d}{ds} \left[\frac{\prod_{j=1}^n (s - p_j)}{\prod_{i=1}^m (s - z_i)} \right] \quad (44)$$

goes to zero. If the locus passes through a region surrounding a saddle point it will be found that the distance along the locus for a given percentage change in gain increases dramatically; furthermore, in plotting the locus the points s_0 lying on the locus will be difficult to find. This situation usually occurs when two separate loci approach each other from opposite directions, when it will be difficult to determine if the locus is as shown (for example) in Figure 4a or 4b. From the engineering point of view, it makes little difference; the location of a root lying in the region of a saddle point will be sensitive to parameter variations, and the system should be designed accordingly, i. e., to be insensitive to the precise location of the root. This figure shows a typical situation encountered when analyzing the low-frequency parasitic modes (represented by the pole-zero dipole near the $j\omega$ axis). If the operating gain is either K_1 or K_2 , the situation is clear. But an intermediate value of gain will result in two complex conjugate pairs of roots within the region of the saddle point that are difficult to locate with precision. This can best be appreciated by one familiar with the use of the Spirule: the locus is hard to "pin down." The difficulty is less severe if the brute force methods of machine computation are used, but even here the number of iterations required to locate the roots will be somewhat larger than usual. It is not at all unusual to run across difficult cases where sophisticated computational techniques are required; the ordinary routines fail to converge.

Another difficulty is in the solution of Equation 43 for the breakaway points. As stated in many textbooks, one can ignore the complex roots if they are far removed from the real axis relative to distances to the nearby poles and zeros on the real axis;

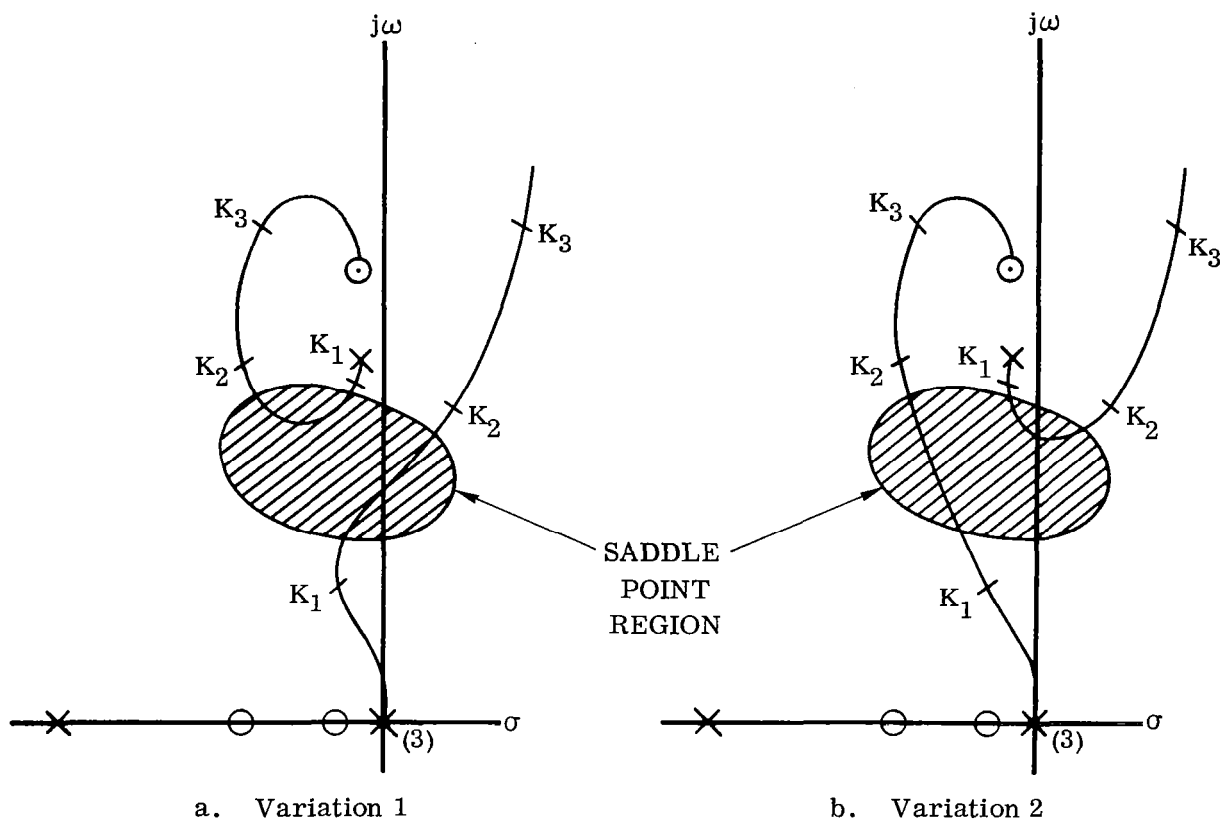


Figure 4. Variations in Root Loci Behavior in the Region of a Saddle Point

neglected are all but the dominant terms in Equation 43. In cases where this is not true, one can use Chen's procedure (Reference 4): from a test point, σ_B , on the real axis one draws a line to one of the pair of complex conjugate poles or zeros being considered. The perpendicular bisector of this line will intersect the real axis at point σ_F . This point is used as an equivalent pole or zero replacing the complex conjugate pair as far as Equation 43 is concerned. Point σ_F changes for every trial value of σ_B . This procedure is involved, but does represent an improvement over direct computations of Equation 43. Fortunately, the precise location of the breakaway point is relatively unimportant in the determination of the behavior of roots near the $j\omega$ axis.

For synthesis purposes, one is concerned with the effect of lead or lag compensation in the open-loop transfer function on the root location or, what amounts to the same thing, on the configuration of the loci. In essence the effect is as follows: in the upper half of the s -plane, loci emanating from complex poles will rotate clockwise about the pole as lag compensation is added. They also will tend to decrease in length (the closed-loop roots move closer to the open-loop poles) by an amount proportional to the attenuation introduced by the lag. The amount of rotation is (approximately) equal to the phase lag introduced at the frequency corresponding to the system root of

interest. In terms of Figure 4, the situation shown in (a) represents somewhat more lag than (b). It also shows the effects of lag on loci emanating from poles on the real axis: the loci will tend to move toward the right-half plane. For lead compensation, the situation is exactly reversed: the loci tend to increase in length and rotate counter-clockwise.

Phase and gain margins are conveniently measured for parasitic mode roots on root locus plots if the definitions in Figure 5 are used. One draws a vector between the pole and the nominal closed-loop root location, and inscribes a circular arc about the open-loop pole of radius equivalent to the distance of the closed-loop root from the pole. A vector is drawn from the pole to the intersection of the arc with the $j\omega$ axis. The angle ϕ_m between the two vectors is the phase margin. The gain margin is the ratio between the gain at the point where the locus crosses the axis and the gain at the operating point (in decibels):

$$\text{Gain Margin} = 20 \log_{10} \frac{K_2}{K_1} \quad (45)$$

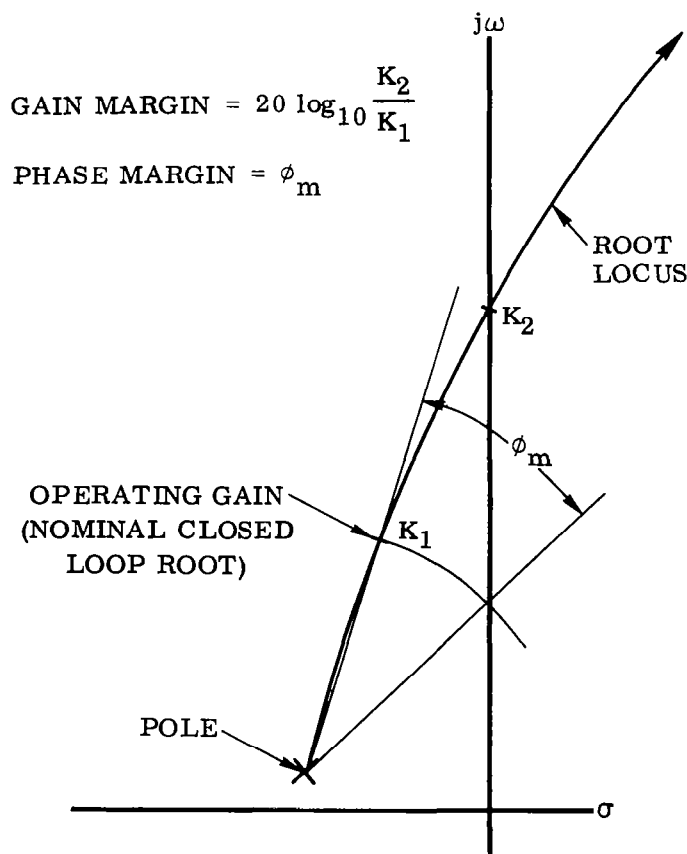


Figure 5. Gain and Phase Margin, Root Locus Interpretation

The gain and phase margins thus measured are analogous, but not equivalent (except approximately) to those measured on a Nyquist diagram (see Paragraph 4.2.3.3, following). If the root locus is for variation of a parameter affecting only the gain of the open-loop transfer function, then the gain margin defined here is exactly equivalent to that measured on the Nyquist diagram. If any parameter is used which affects the open-loop phase-versus-frequency characteristic of $GH(s)$ then this will no longer be true. The phase margin as defined here is equivalent to that measured on the Nyquist diagram to the degree that the open-loop pole is far removed from other poles and zeros (the pole has a very small residue). In this case the root locus approximates a straight radial line from the pole.

For exact equivalence between gain the phase margins measured on root loci and Nyquist plots one must plot constant gain and constant phase loci in the s -plane. This is equivalent to finding those points, s_0 , satisfying Equation 37 for fixed K with no regard to the conditions imposed by Equation 38 (constant gain loci) and similarly, finding points s_0 satisfying Equation 38 with an arbitrary phase angle ϕ (positive or negative) added to the right-hand side of this equation. If these plots are made throughout the s -plane, the result is exactly analogous to stream lines (constant phase loci) and equipotential lines (constant gain loci) in a potential problem. In the region of an isolated pole (i. e., near enough to the pole such that the other poles and zeros contribute little to the evaluation of Equations 37 and 38) the appearance is that of radial lines (constant phase loci) emanating from the pole and circles (constant gain loci) surrounding the pole, the two crossing each other at right angles. When other poles or zeros are nearby, the plot is distorted from this ideal form, but it remains orthogonal. The gain and phase margins can be read off by inspection and used in the determination of the required flight control system tolerances. Further discussion on this point (margins) is deferred to a succeeding section where the stability margins are defined on the Nyquist diagram.

The procedure of the preceding paragraph amounts to introducing a complex gain function, $Ke^{j\phi}$, within the closed loop. Such a function has no meaning in reality; it is convenient only in the mathematical sense. It is sometimes more pertinent to determine the effects of physically realizable phase and gain changes within the closed loop. The procedure here is clear: one introduces appropriate poles and zeros of the element introduced (a lead-lag filter for example) and plots the loci. In practice the open-loop gain, poles, and zeros representing the autopilot, sensors, and actuation are perturbed about their nominal values throughout their range of variation (based on parameter tolerances) and the effect on the root locations noted. From the opposite point of view, this gives an indication of required tolerances on flight control system parameters.

4.2.2.4 Time and Frequency Response. The root locus technique is fundamentally a means of locating the roots of the characteristic equation. Its chief virtue is the display of these roots as functions of various system parameters. But most important are the system response properties to command or disturbance inputs, both inferred from the location of the closed-loop system poles and zeros. The s -plane diagram of

these poles and zeros serves as a graphical means for determining the time and frequency response of the system.

The transfer function of a system can be represented by

$$T(s) = \frac{K \prod_{i=1}^m (s - z_i)}{\prod_{j=1}^n (s - p_j)} \quad (46)$$

The input can be a command, a disturbance, or a test input; the output is the variable whose response is of interest. Consider the frequency response obtained by substituting $j\omega$ for s ,

$$\begin{aligned} T(j\omega) &= \frac{K \prod_{i=1}^m (j\omega - z_i)}{\prod_{j=1}^n (j\omega - p_j)} \\ &= |T(j\omega)| e^{j \text{Arg } T(j\omega)} \end{aligned} \quad (47)$$

This may be evaluated graphically from the s -plane plot of the poles and zeros for any desired ω by picking a point on the positive $j\omega$ axis corresponding to the desired frequency and measuring off distances to all the poles and zeros and their respective phase angles according to

$$\begin{aligned} |T(j\omega)| &= \frac{K \prod_{i=1}^m |j\omega - z_i|}{\prod_{j=1}^n |j\omega - p_j|} \\ \text{Arg } T(j\omega) &= \sum_{i=1}^m \text{Arg}(j\omega - z_i) - \sum_{j=1}^n \text{Arg}(j\omega - p_j) \end{aligned} \quad (48)$$

where the angles are measured at the poles and zeros counterclockwise from the 3 o'clock position to the $j\omega$ point on the $j\omega$ axis. This can be done using the Spirule.

Consider the transform of the time response to an input, $C(s)$.

$$C(s) = R(s) T(s)$$

$$= \frac{K \prod_{i=1}^m (s - z_i)}{\prod_{j=1}^n (s - p_j)} R(s) \quad (49)$$

If $R(s)$ is in closed form having only first-order poles (this excludes ramp inputs, for example) at frequencies different from the system poles, the response may be evaluated from an s -plane plot of the poles and zeros of $C(s)$. $C(s)$ may then be represented by:

$$C(s) = \frac{K' \prod_{i=1}^k (s - z_i)}{\prod_{j=1}^{\ell} (s - p_j)} \quad (50)$$

where the indices k and ℓ (rather than m and n) indicate additional poles and zeros in $C(s)$ due to $R(s)$. $C(s)$ has the partial fraction expansion,

$$C(s) = \sum_{j=1}^{\ell} \frac{A_j}{(s - p_j)} \quad (51)$$

where the p^{th} residue is

$$\begin{aligned} A_p &= (s - p_p) C(s) \Big|_{s=p_p} \\ &= |A_p| e^{j\phi} \end{aligned} \quad (52)$$

That is,

$$\left. \begin{aligned} |A_p| &= K' \prod_{i=1}^k |p_p - z_i| / \prod_{\substack{j=1 \\ j \neq p}}^{\ell} |p_p - p_j| \\ \phi &= \sum_{i=1}^k \text{Arg}(p_p - z_i) - \sum_{\substack{j=1 \\ j \neq p}}^{\ell} \text{Arg}(p_p - p_j) \end{aligned} \right\} \quad (53)$$

In short, the residue at the p^{th} pole of $C(s)$ is determined by measuring the distance and angles as before from the s -plane plot. The time response is given by

$$C(t) = \sum_{j=1}^{\ell} A_j e^{p_j t} \quad (54)$$

If some of the poles are complex, the residues appear in complex conjugate pairs. Suppose there are poles at $p_1 = (\sigma_1 + j\omega_1)$ and $p_2 = (\sigma_1 - j\omega_1)$ with residues $A_1 e^{j\phi}$ and $A_1 e^{-j\phi}$. The corresponding contribution to the time response $C(t)$ is

$$C_1(t) = 2 A_1 e^{\sigma_1 t} \cos(\omega_1 t + \phi) \quad (55)$$

The corresponding time functions may be graphically added to give the overall time response.

4.2.2.5 Example - Rigid Body Flight Control. Consider the system under discussion in the preceding subsection. The open-loop transfer function is given by

$$GH(s) = \frac{K_A \mu_{\xi} (s + K_I)(K_R s + 1)}{s(s^2 - \mu_{\beta})(\tau_c s + 1)} \quad (56)$$

or in the form of Equation 36,

$$GH(s) = \frac{K_A \mu_{\xi} K_R}{\tau_c} \frac{(s + K_I) \left(s + \frac{1}{K_R} \right)}{s(s^2 - \mu_{\beta}) \left(s + \frac{1}{\tau_c} \right)} \quad (57)$$

For the system parameters of Tables 1 and 2, the root loci are given by Figure 6 in terms of K_A . Note that for display purposes only the upper half of the s -plane is necessary, the loci being symmetric with respect to the real axis. But suppose it is desired to determine the variations in the system roots with K_R . The characteristic equation may be manipulated to give

$$1 + \frac{K_A \mu_{\xi} K_R}{\tau_c} \frac{s(s + K_I)}{s(s^2 - \mu_{\beta}) \left(s + \frac{1}{\tau_c} \right) + \frac{K_A \mu_{\xi}}{\tau_c} (s + K_I)} = 0 \quad (58)$$

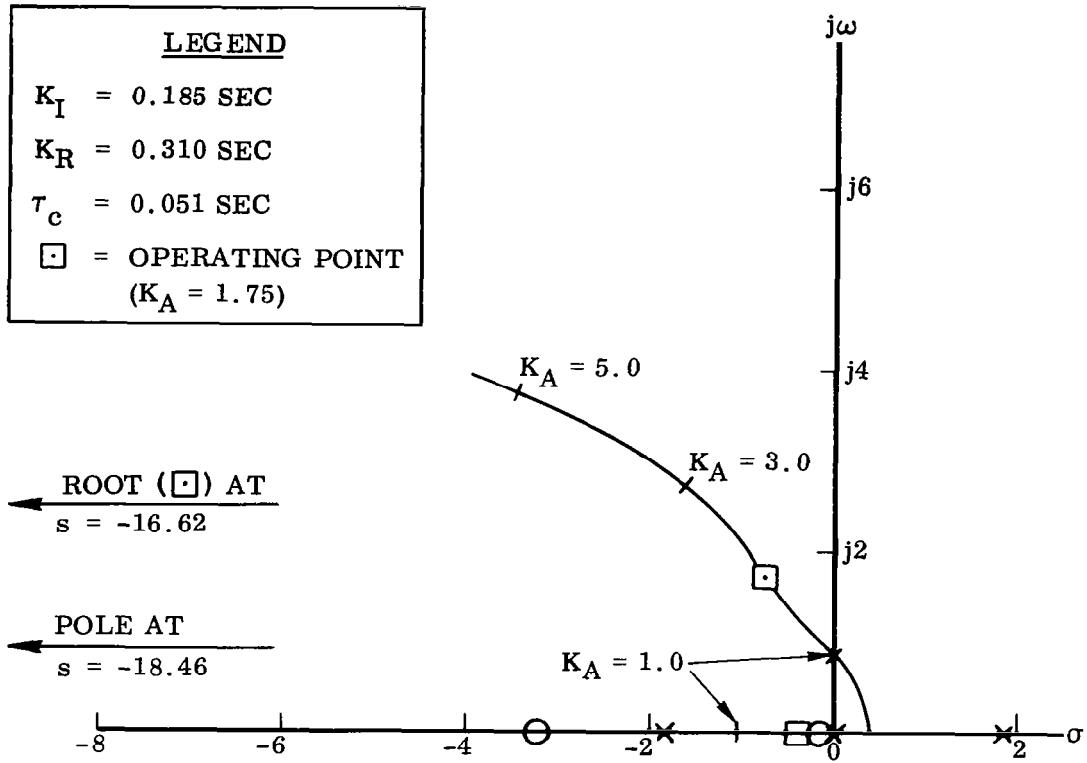


Figure 6. Rigid Body Root Locus Under Forward Loop Gain Variation

where the poles of the denominator may be found by determining root locations of

$$1 + \frac{K_A \mu_\xi}{\tau_c} \frac{s + K_I}{s(s^2 - \mu\beta)\left(s + \frac{1}{\tau_c}\right)} = 0 \quad (59)$$

which is recognized as the characteristic equation of a system with no rate compensation. This can be done by means of root locus techniques. The loci for variations in K_R are sketched in Figure 7. The effects of increasing rate gain on the root location are clear, and it is seen that open-loop gain need not be the only parameter varied when plotting root loci.

4.2.3 FREQUENCY RESPONSE METHODS. Most of the analysis and synthesis of control systems of all types makes use of one or more of the various frequency response techniques. The method is older than that of root locus and consequently is more familiar to most control engineers. Its basic advantage is that of synthesis from knowledge of the open-loop frequency response (steady-state response to an input sinusoid of fixed frequency as a function of frequency) without necessarily knowing precise pole and zero locations of the transfer functions. This enables one to use experimental

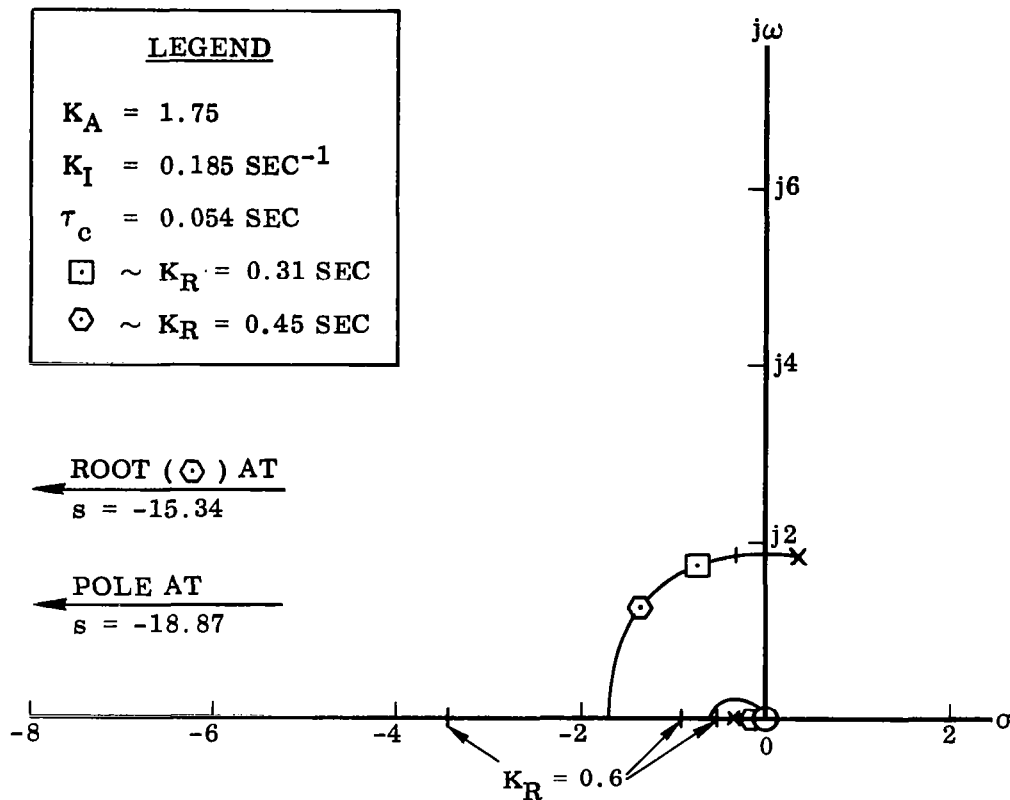


Figure 7. Rigid Body Root Locus Under Rate Gain Variation

frequency response data directly in the analysis and synthesis of the booster flight control systems. Another advantage from the computational standpoint is that it is unnecessary to factor the polynomial representing the characteristic equation; the relative stability of the closed-loop system may be inferred from the open-loop response. Its major disadvantage is difficulty of interpretation as the system roots are not explicitly available for determination of the system time response. The following discussion is intended to outline basic features of this approach, not to function as a rigorous theoretical discussion found in textbooks.

4.2.3.1 Definitions. The characteristic equation of a linear control system may be written in the form,

$$1 + GH(s) = 1 + K \frac{\prod_{i=1}^m (\tau_i s + 1)}{\prod_{j=1}^n (\tau_j s + 1)} \quad (60)$$

where τ_i and τ_j are the system time constants (both real and complex) and K is the "low-frequency open-loop gain" of the system. As before, $GH(s)$ is the open-loop

system transfer function, a function of s , the Laplace transform variable. The frequency response approach to solution of the synthesis problem consists of shaping $GH(j\omega)$ such that Equation 60 cannot be satisfied for any value of complex frequency s having a positive real part. This is done by comparing $GH(j\omega)$ against the Nyquist criterion (Paragraph 4.2.3.2) by one of a number of methods of plotting $GH(j\omega)$.

4.2.3.1.1 Bode Plots. The most familiar form is the Bode plot of the gain $|GH(j\omega)|$ in decibels (given by $20 \log_{10} |GH(j\omega)|$) and the phase $\text{Arg } GH(j\omega)$ in degrees versus a logarithmic frequency scale (usually in cycles per second but sometimes in radians per second). The utility of these plots arises from two major features:

1. Elements added in cascade with $GH(j\omega)$ are represented by algebraic additions of gain (in decibels) and phase (in degrees).
2. The gain of cascaded elements may be approximated by straight lines (except in the case of underdamped poles and zeros) as functions of frequency, the actual responses differing from the approximations by known amounts.

In practice the control engineer has access to standard plots of gain and phase versus normalized frequency of the standard elements, first- or second-order lags or leads. The frequency is normalized with respect to the break frequency of first-order elements (the frequency of the pole or zero) and the undamped natural frequency of second-order elements (where the damping ratio is less than unity). Synthesis via Bode plots can therefore be done graphically on semi-log graph paper using a straight edge.

4.2.3.1.2 Nichols Charts. A second form in wide use is Nichol's chart. Here the ordinate is the gain in decibels and the abscissa is phase angle in degrees. $GH(j\omega)$ becomes a curved line with frequency as a parameter. Here again, frequency characteristics of the individual factors of the transfer function may be added graphically. Nichol's chart is sometimes referred to as the W -plane by virtue of the following transformation of variables.

$$\left. \begin{aligned} W(j\omega) &= \ln GH(j\omega) \quad (\text{principal value}) \\ &= \ln C(\omega) + j\phi(\omega) \end{aligned} \right\} \quad (61)$$

The open-loop gain is in nepers plotted on the real axis, and the phase in radians on the imaginary axis of the W -plane. These units are inconvenient (not to say unfamiliar) and a change in scale is accomplished as follows.

$$\left. \begin{aligned} W' &= (20 \log_{10} e) \text{Re } W + j \frac{180}{\pi} \text{Im } W \\ &\cong 8.68 \text{Re } W + j 57.3 \text{Im } W \end{aligned} \right\} \quad (62)$$

which is Nichol's chart. Note that mapping is no longer conformal unless the scales are chosen properly; the unit of length representing one decibel must be equivalent to 6.6 degrees in phase. This is not usually the case. In practice, a choice of, say, one inch equaling 10 db and 90 degrees will result in a relative error between the scales of less than five percent.

4.2.3.1.3 Nyquist Diagrams. A third form is the polar plot of $GH(j\omega)$, known as the Nyquist diagram. The open-loop gain $GH(j\omega)$ is represented by a radial vector from the origin whose length is $|GH(j\omega)|$ and whose angle (measured counterclockwise with respect to the positive real axis, the three o'clock position) is $\text{Arg } GH(j\omega)$. The frequency characteristics of various elements of the transfer function may no longer be added; the magnitudes must be multiplied and phase angles added. As in Nichol's chart, frequency in radians/second is the parameter of the curve. Nichol's chart is seen to be a one-to-one mapping of the Nyquist diagram and the two will be treated as one in the discussions that follow.

4.2.3.1.4 Mapping. A central concept lying at the heart of frequency response methods is that of mapping. As indicated above, Nichol's chart and Nyquist diagram are maps of each other and both are maps of the Bode diagram. Further, the GH-plane on which the Nyquist diagram is plotted is mappable from the s-plane on which poles and zeros of the transfer function $GH(s)$ are plotted. Any point in the s-plane is mapped into the GH-plane under the transformation $GH(s)$. Mapping is one-to-one in this direction, i.e., each value for s corresponds to a single point in the GH-plane. The converse is not true. A point in the GH-plane corresponds to several points in the s-plane; in particular, n points where n is the number of poles of $GH(s)$. (It is assumed that $n \geq m$, the number of zeros of $GH(s)$.) This is a consequence of the fundamental theorem of algebra that states that the n^{th} -order equation in s (the characteristic equation of the control system) has n solutions. A final property of importance is the conformal nature of the mapping. A small region in the s-plane maps into a similarly small region in the GH-plane. This holds true for every point in the s-plane except where $GH(s)$ has poles, zeros, or saddle points. If the small region in the s-plane encloses such points there will be considerable distortion in mapping from s-plane to GH-plane. In the above context, the word "small" means a region not containing a pole, small enough that $GH(s)$ does not vary appreciably in the region; that is, small compared with the distances between poles and zeros in the s-plane. Further discussion on mapping may be found in most mathematical texts on complex variable theory; for example, Reference 5.

4.2.3.2 Nyquist Criterion. Application of the Nyquist criterion affords a means of interpreting the stability qualities of the closed-loop control system from the open-loop frequency response $GH(j\omega)$ plotted in the GH-plane. The theorem stating this result is a direct consequence of Cauchy's Residue Theorem in complex variable theory. This theorem states that the line integral around a closed curve C of a real function of a complex variable s in the s-plane is equal to $2\pi j$ ($j = \sqrt{-1}$) times the sum of the residues at the poles within the region enclosed by C . Mathematically, this is

$$\int_C f(s) ds = 2\pi j \left(\sum_{i=1}^k b_i \right) \quad (63)$$

where C encircles k poles, $p_1, p_2 \dots p_k$, of $f(s)$ and where $f(s)$ has the partial fraction expansion,

$$f(s) = \sum_{i=1}^k \frac{b_i}{(s - p_i)} + (\text{additional terms due to multiple order poles}) \quad (64)$$

From this it follows that for complex variable GH ,

$$\int_D \frac{d(GH)}{(GH)} = 2\pi j N \quad (65)$$

where N is the number of times the closed curve D encircles the origin in the GH -plane. (The residue of $(GH)^{-1}$ is 1 and the pole is at the origin in the GH -plane.) It is now necessary to relate Equations 63 and 65.

Suppose $GH(s)$ has a zero, s_0 , of multiplicity α . A Taylor's series expansion about this point results in

$$GH(s) = 1 + a_\alpha (s - s_0)^\alpha + a_{\alpha+1} (s - s_0)^{\alpha+1} + \dots \quad (66)$$

The first derivative at this point is given by

$$GH'(s) = \alpha a_\alpha (s - s_0)^{\alpha-1} + (\alpha + 1) a_{\alpha+1} (s - s_0)^\alpha + \dots \quad (67)$$

Equation 67 divided by Equation 66 results in

$$\frac{GH'}{GH}(s) = \alpha (s - s_0)^{-1} + c_0 + c_1 (s - s_0) + \dots \quad (68)$$

This implies a pole of first order at s_0 having residue α , the multiplicity of the zero at s_0 in the function $GH(s)$. Now suppose $GH(s)$ to have a pole at s_p of multiplicity β . A Taylor's series expansion about this point results in

$$GH(s) = b_{-\beta} (s - s_p)^{-\beta} + b_{-\beta+1} (s - s_p)^{-\beta+1} + \dots \quad (69)$$

The first derivative is

$$GH'(s) = -\beta b_{-\beta} (s - s_p)^{-\beta-1} + (-\beta + 1) b_{-\beta+1} (s - s_p)^{-\beta} + \dots \quad (70)$$

Finally,

$$\frac{GH'}{GH}(s) = -\beta (s - s_p)^{-1} + d_0 + d_1 (s - s_p) + \dots \quad (71)$$

Thus the function $\frac{GH'(s)}{GH(s)}$ has a first-order pole at s_p of residue $-\beta$, the multiplicity of the pole in the function $GH(s)$.

Now the line integral about a closed curve C encircling k poles of a function $f(s)$ is equivalent to a summation of the line integrals about small closed curves c_i about each of the poles p_i ($i = 1, 2, \dots, k$) enclosed by C .

$$\int_C f(s) ds = \sum_{i=1}^k \int_{c_i} f(s) ds \quad (72)$$

Hence it follows that

$$\frac{GH'(s) ds}{GH(s)} = 2\pi j \left(\sum_{i=1}^{\ell} \alpha_i - \sum_{i=1}^g \beta_i \right) \quad (73)$$

where C encloses ℓ zeros of multiplicity α_i ($i = 1, 2, \dots, \ell$) and g poles of multiplicity β_i ($i = 1, 2, \dots, g$). But

$$\int_C \frac{GH'(s) ds}{GH(s)} = \int_D \frac{d(GH)}{(GH)} \quad (74)$$

where the closed curve D in the GH -plane is a mapping of the curve C in the s -plane.

It is now a simple matter to transfer interest from the open-loop case to the closed-loop case, whence

$$\int_C \frac{GH'(s) ds}{1 + GH(s)} = \int_D \frac{d(GH)}{(1 + GH)} \quad (75)$$

where D is unchanged but the number of encirclements of the $(-1 + j0)$ point in the GH -plane is of interest:

$$N = N_z - N_p \quad (76)$$

where N_z is the number of zeros and N_p the number of poles (both counted as to multiplicity) of the function $1 + GH(s)$ enclosed by curve C in the s -plane. The zeros are the roots of the characteristic equation of the system. The poles are the poles of open-loop transfer function $GH(s)$. It remains only to choose curve C to enclose the region of the s -plane resulting in instability, i.e., the right-half plane. This is done by choosing C to be the $j\omega$ axis in the s -plane from $-jR$ to $+jR$, excluding poles that appear on the $j\omega$ axis by small semicircles of radius r . The curve is closed to the right by a large semicircle of radius R centered at the origin. One then allows R to approach infinity and r to approach zero, and curve C surrounds the entire right-half plane. A sketch of this construction is shown in Figure 8. Note that the curve is traversed in a clockwise direction.

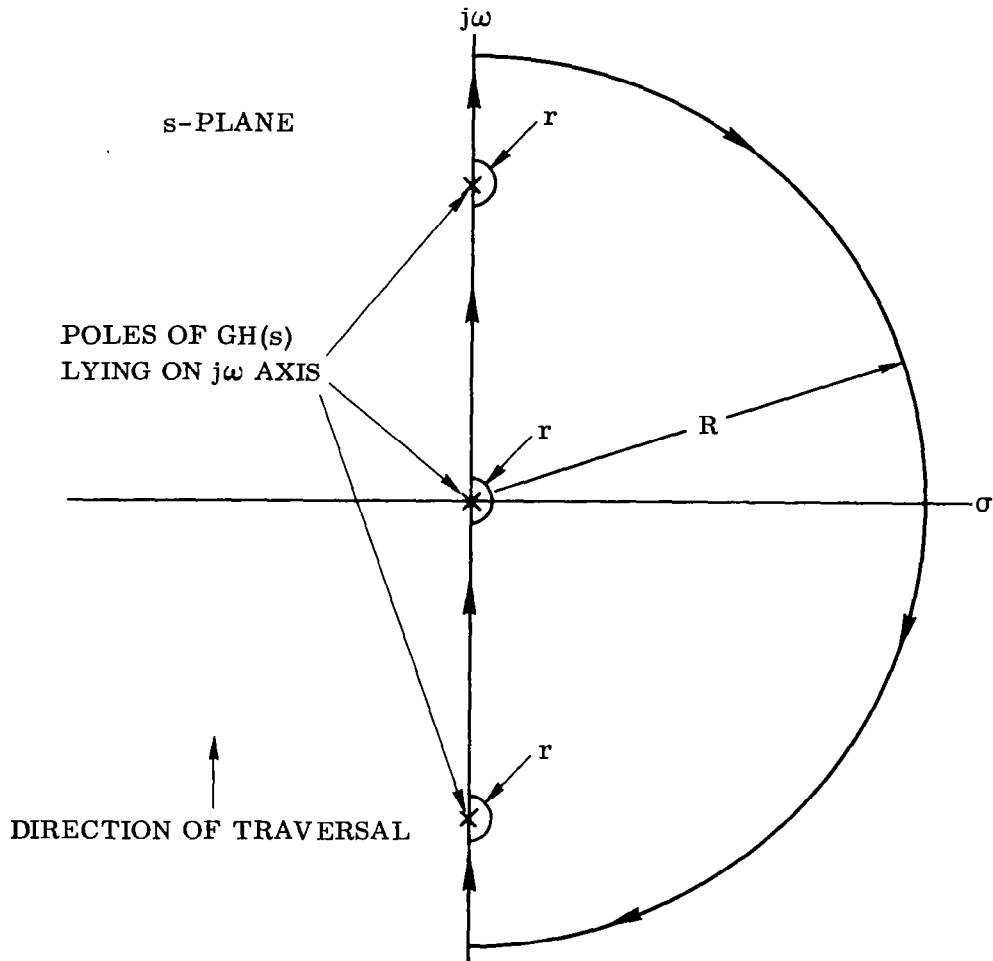


Figure 8. The Curve C Used as the Nyquist Criterion for Stability

To map this curve onto the GH-plane, one plots function $GH(j\omega)$ in the GH-plane as a function of ω ($-\infty \leq \omega \leq +\infty$). As ω approaches a pole on the $j\omega$ axis, $GH(j\omega)$ becomes very large; curve D in the GH-plane approaches infinity. As the semicircle bypassing a simple pole is traversed, a phase shift of approximately 180 degrees takes place. This corresponds to a semicircle near infinity on the GH-plane. Because the mapping is conformal the semicircle goes in a clockwise direction. This can most easily be remembered by associating the small right angle at the start of the semicircle with the right angle D must make near infinity. A similar argument holds when the semicircle is complete; again D must turn to the right. If the small semicircle bypasses a double or triple pole (found at the origin in flight control systems) the phase shift is 360 or 540 degrees and the curve D on the GH-plane makes one or one-and-one-half circuits near infinity. The result of this procedure is a rather complex curve (except for simple systems) on the GH-plane. The Nyquist criterion can then be stated as follows:

"The net number of clockwise encirclements of the point $-1 + j0$ in the GH-plane plus the number of poles (counted as to multiplicity) of $GH(s)$ lying in the right-half s -plane (determined by inspection) is equal to the number of roots of the closed-loop system lying in the right-half s -plane."

It is important to note the sense of D. A counterclockwise encirclement results in $N = -1$, for example.

In practice, the Nyquist diagram that results from this procedure is plotted only for positive values of ω ; that is, a polar plot of the open-loop frequency response. The resultant curve must be completed (at least mentally) to determine N, the number of clockwise encirclements. This is easily done when it is remembered that the curve for negative ω is a mirror image of the curve for positive ω about the real axis in the GH-plane. The semicircles near infinity must also be supplied. This is done by careful observation of the direction of the frequency change along the curve as it approaches infinity and knowledge of the placement and multiplicity of the open-loop poles on the $j\omega$ axis.

4.2.3.3 Interpretation of Nyquist Diagrams. The Nyquist criterion as stated above is similar to the Routh-Hurwitz criterion; the question of stability has a yes or no answer. The importance, however, is that it affords a means of estimating the relative system stability by observing the nearness of curve D to the -1 point in the GH-plane. This point indicates where the characteristic equation is satisfied; hence it is the point where all closed-loop roots are mapped into the GH-plane. Curve D is a mapping of the $j\omega$ -axis. The nearness of approach is therefore a measure of the nearness of one or more system roots to the $j\omega$ axis. In fact, this idea can sometimes be used to estimate a root location when the root has a small real part.

4.2.3.3.1 Estimation of Roots Having Small Real Parts. Consider the sketches of Figure 9. The first sketch plots the usual Nyquist plot ($\sigma = 0$) of an example system on

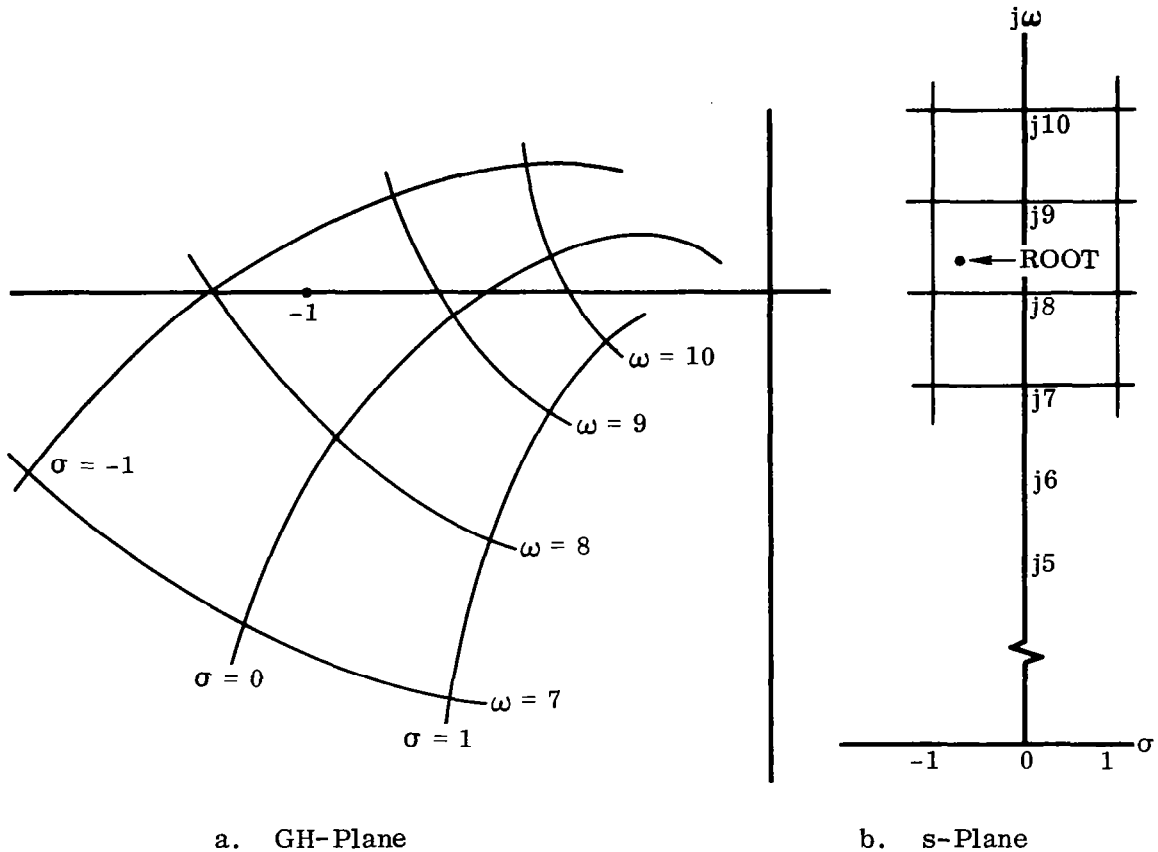


Figure 9. Root Location on Nyquist Diagram

the GH-plane together with mapping of the s-plane grid in the immediate vicinity of the root. The second sketch shows the region of s-plane being mapped. The mapping is conformal and the result on the GH-plane is a set of curvilinear coordinates. In practice, only the Nyquist plot need be used as the orthogonal net can be sketched in for a short distance away from the curve. In this example, the root is located at $-0.7 + j8.4$ and therefore has a per-unit critical damping of 0.083. The same result can be obtained by measuring the shortest distance to the Nyquist curve in units of frequency, the scale being that of the frequency gradient along the curve adjacent to the minus-one point. This distance is the real part of the root; the point where the perpendicular intersects the curve is the imaginary part. This technique is quite useful for estimating the location of parasitic mode roots from Nyquist plots. Unfortunately, it sometimes fails, as for example when the point of closest approach is not clearly defined, or where there is a nearby saddle point in the s-plane resulting in heavy local distortion of the orthogonal net.

4.2.3.3.2 Gain and Phase Margins. A further observation is that the negative real axis in the GH-plane is a mapping of the root loci under gain variation. If the open-loop gain is increased or decreased, the Nyquist plot expands or contracts; any point on

D moves along radial vectors from the origin. Rather than replotting the Nyquist curve, the scale can be varied inversely with the open-loop gain. Thus the minus-one point moves along the negative real axis of the plot. Motion toward the origin corresponds to increasing gain, and conversely for motion away from the origin.

Rotation of the axes of the Nyquist plot about the origin relative to the plotted curve corresponds to multiplication of the open-loop gain by $e^{j\phi}$, where ϕ (phase) is the angle of rotation. With respect to the mapping of the positive $j\omega$ axis, a clockwise rotation represents increasing lead ($\phi > 0$) and conversely for counterclockwise rotation.

This leads into the concepts of phase margin and gain margin. The gain margin is defined as that change in the gain of a stable system with no change in phase required to result in marginal stability (the Nyquist curve passing through the minus-one point). Phase margin is similarly defined as the phase change required (with no change in gain) to destabilize the nominally stable system. In terms of Figure 10, if the gain is reduced by a multiplicative factor of $\frac{1}{a}$, the system goes unstable because of too low a gain. The quantity $\frac{1}{a}$ (usually quoted in decibels, i.e., $20 \log_{10} \frac{1}{a}$) is the negative gain margin for the system. A similar definition holds for positive gain margin, the gain ($\frac{1}{b}$ in the figure) increase required for marginal stability. The phase margin is defined for this figure as the amount of additional lag with no change in gain required for marginal stability.

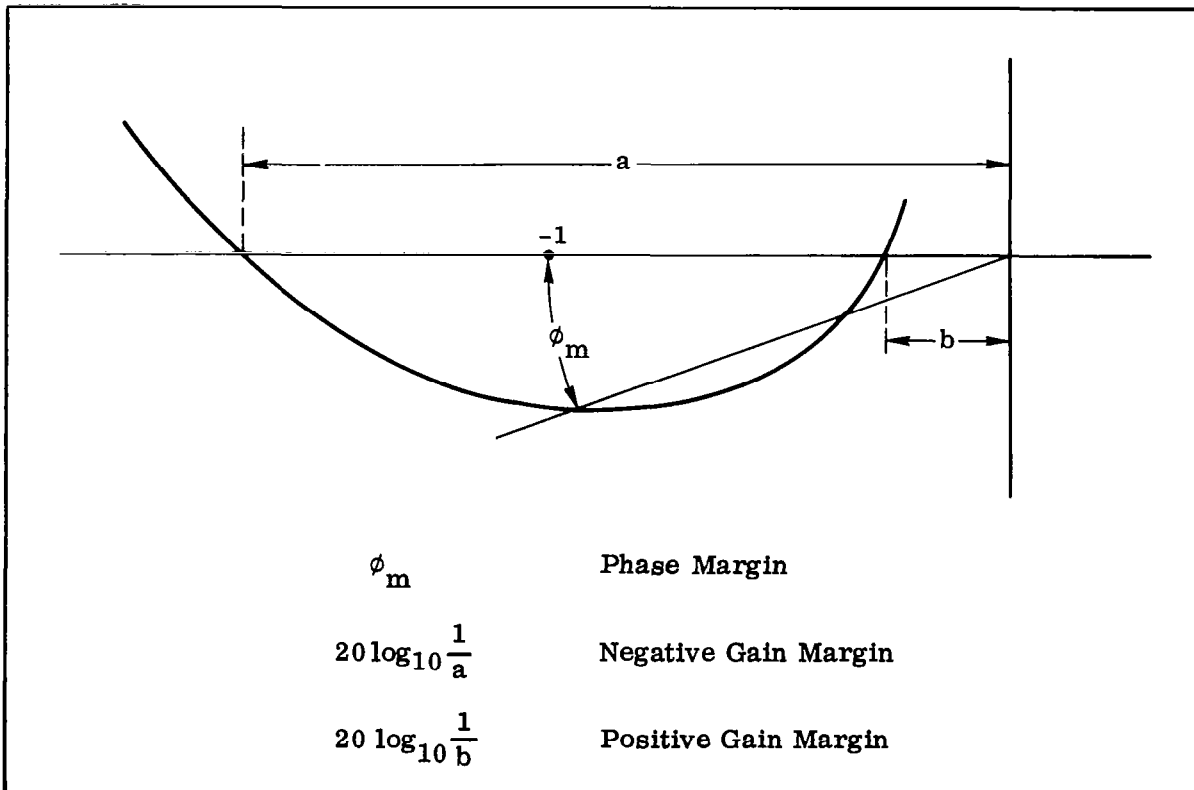


Figure 10. Nyquist Diagram for Conditionally Stable System, Showing Gain and Phase Margin

4.2.3.3.3 Complex Gain Margin - Zone of Exclusion. These concepts of gain and phase margin are not always adequate to ensure satisfactory margin for stability in launch vehicle flight control system design. If certain gain and phase margins are specified, a "zone of exclusion" is defined in the GH-plane consisting of a portion of the negative real axis and an arc drawn through the minus-one point. If the Nyquist plot crosses neither of these lines, the system has sufficient margin under this definition. Consider the sketch of Figure 11. Both gain and phase margins for the lightly damped root indicated are infinite, yet serious question exists as to system adequacy under parameter variations of any size. Since this situation is sometimes encountered with parasitic modes of response, it would be desirable to define a zone of exclusion which would reject situations such as shown in Figure 11.

Fundamentally, such a zone of exclusion is equivalent to specifying a domain of gain and phase variations allowable in the open-loop transfer function that will ensure satisfactory stability of the system roots. Since a detailed model of the launch vehicle dynamics results in many roots located near the $j\omega$ axis, each corresponding to a particular (usually parasitic) mode of response, it is reasonable to require a different domain of such variations for each root, depending on the effects that likely parameter variations will have on each root and the degree of stability required in the response.

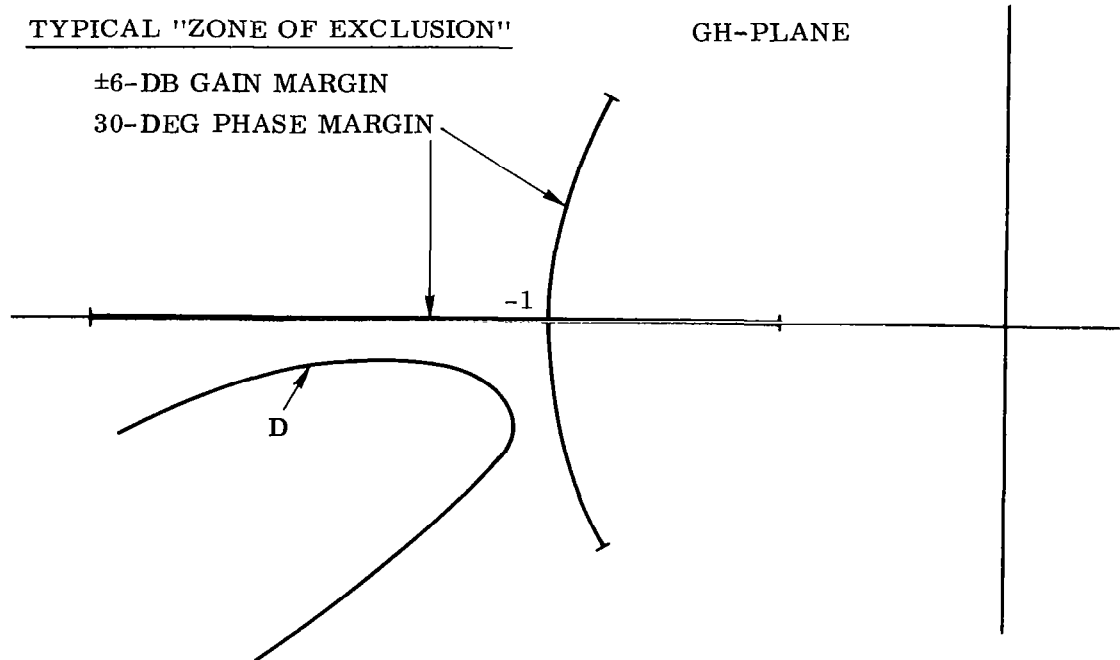


Figure 11. Failure of Gain and Phase Margins, as Usually Defined, to Verify System Adequacy

The question of what constitutes the likely range of parameter variations will be taken up later. For the present it is sufficient to indicate the concept involved without specifying the dimensions of the zone of exclusion desired. The specified domain of gain and phase variations will be called the complex gain margin, given by

$$\text{Complex Gain Margin} = Ke^{j\phi} \quad (77)$$

where K is the magnitude and ϕ the phase of an idealized element in the nominal open-loop transfer function representing the desired margin for stability. K and ϕ are functionally related to form a closed curve in the complex K -plane. Figure 12a illustrates a complex gain margin where increasing gain tends to correlate with increasing phase lead, a typical situation. Figure 12b illustrates the zone of exclusion in the GH-plane, given by inverting $Ke^{j\phi}$ with respect to the minus-one point, i.e.,

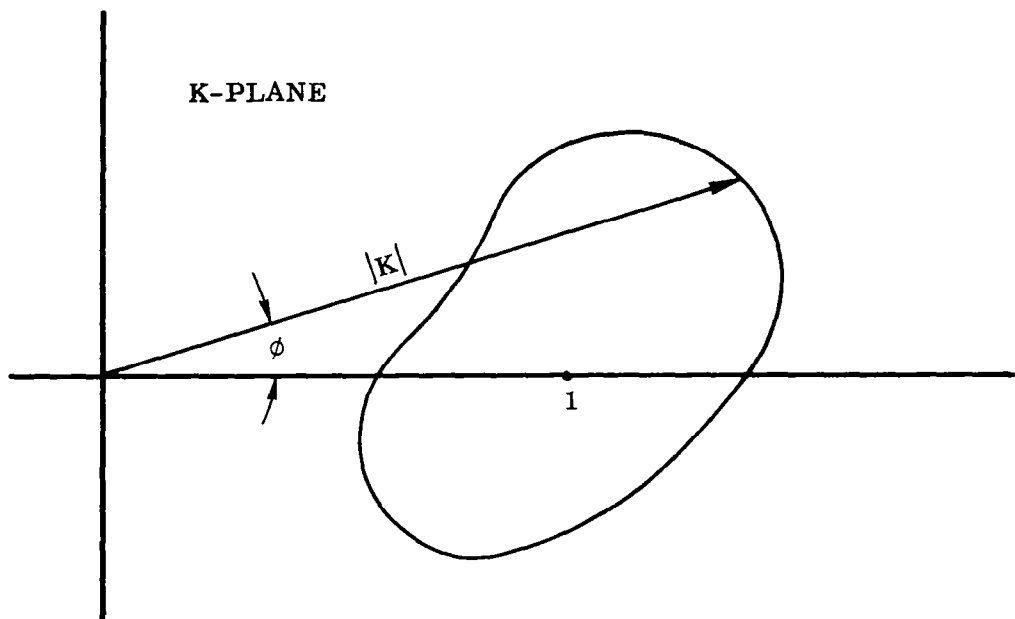
$$\text{Zone of Exclusion} = -\frac{1}{K}e^{-j\phi} \quad (78)$$

For this figure, the inversion was accomplished graphically by reading off K as a function of ϕ from Figure 12a, inverting K , and plotting versus $180^\circ - \phi$ in the GH-plane. The zone of exclusion thus plotted represents that portion of the GH-plane where the Nyquist curve for positive ω is not permitted, and represents a locus of extreme locations for the minus-one point with respect to a fixed plot of the Nyquist curve for positive ω and nominal gain ($Ke^{j\phi} = 1$). In the example this zone of exclusion is heavily weighted against lead, as expected from the choice of $Ke^{j\phi}$.

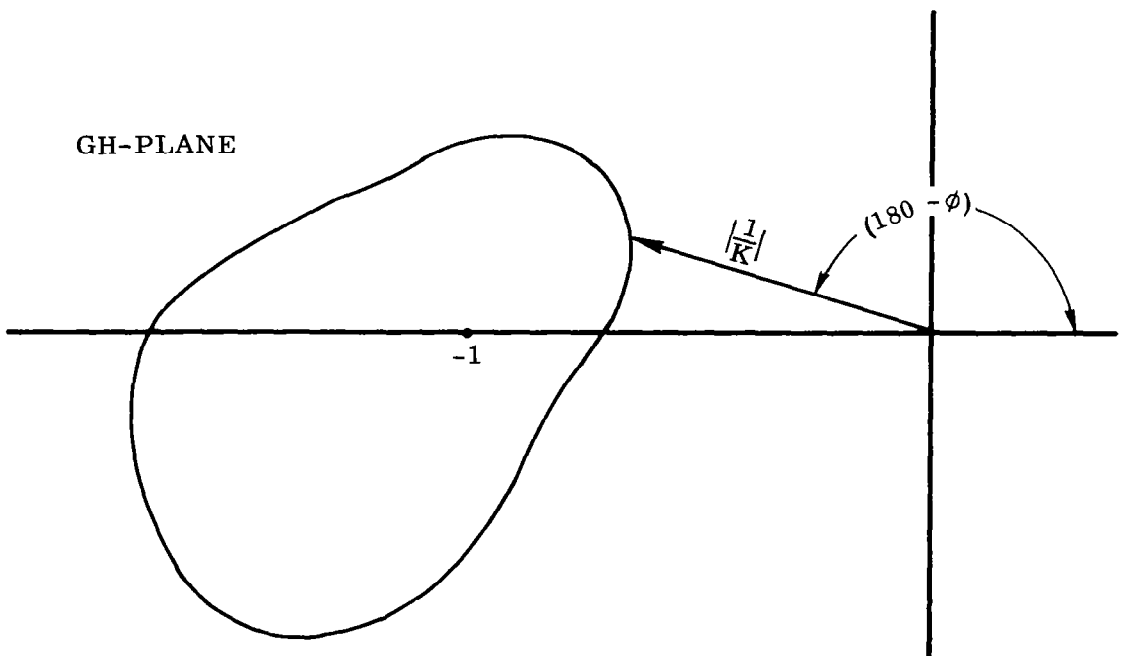
More typically, the flight control system specifications might call for a gain margin of ± 6 db and a phase margin of ± 30 degrees, with no correlation of gain and phase expressed. On the assumption that in the physical situation the gain in decibels and the phase in degrees are both normally distributed statistical variables, a bi-variate distribution in simplified, the extremes of which coincide with the specifications. The complex gain margin then satisfies the equation,

$$\left[\frac{20 \log_{10} |K|}{6} \right]^2 + \left[\frac{\phi}{30} \right]^2 = 1 \quad (79)$$

which is recognized as an ellipse when gain (db) is plotted against phase (deg). See Figure 13a. This interpretation is largely intuitive, the philosophy being that it is unlikely that the physical situation would result in simultaneous extreme variations in both gain and phase (this would correspond to a sector in the complex K -plane whose extremes are defined by the specifications). The zone of exclusion is plotted in Figure 13b. This particular choice of complex gain margin has a plot in the complex K -plane identical to the zone of exclusion but reflected about the imaginary axis. It is interesting to note that the zone of exclusion in this case is very nearly approximated by an

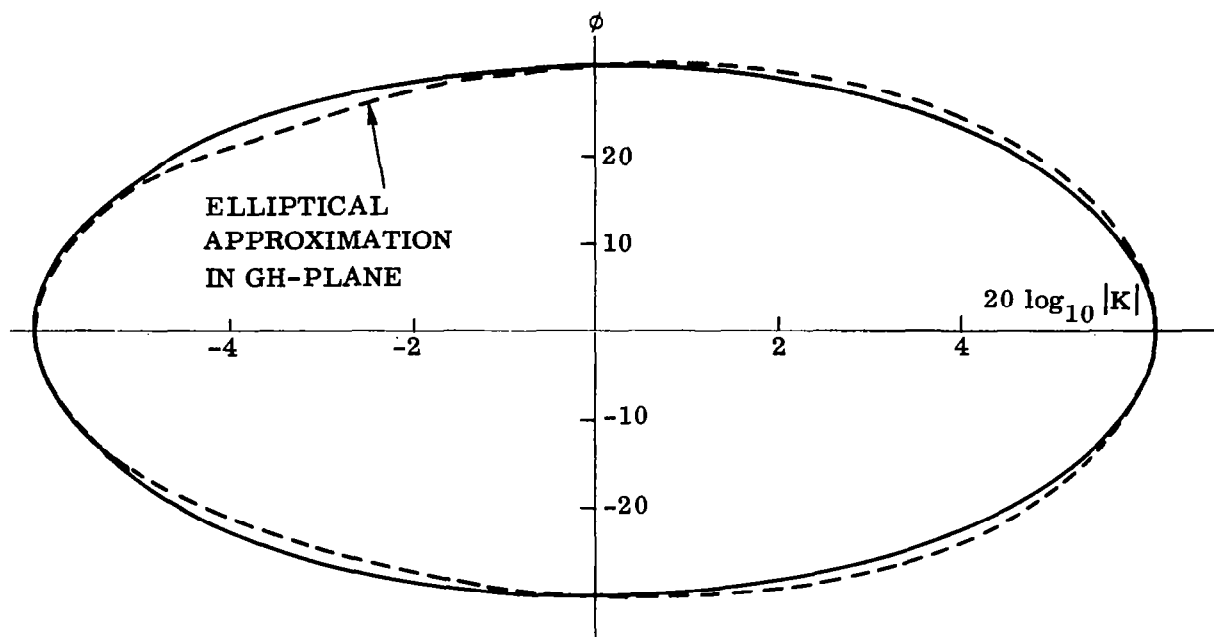


a. Complex Gain Margin

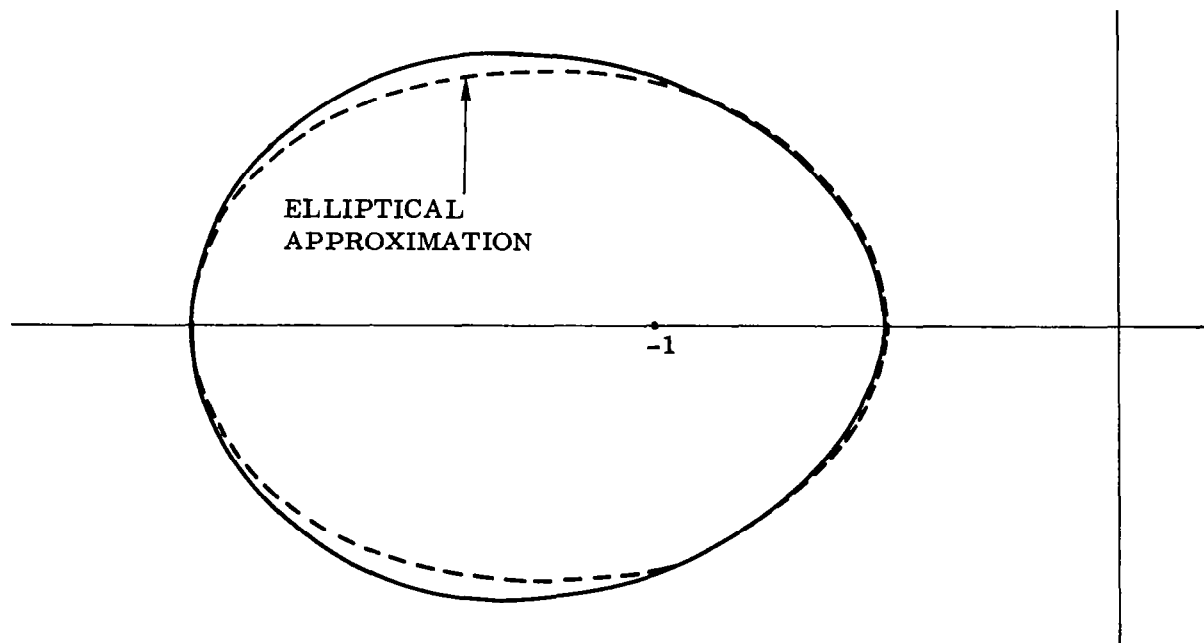


b. Corresponding Zone of Exclusion

Figure 12. Complex Gain Margin



a. Gain (db) Versus Phase (deg)



b. Zone of Exclusion in GH-Plane

Figure 13. Zone of Exclusion for $Ke^{j\phi}$ Having Elliptical Form

ellipse constrained such that the gain margins for zero phase are ± 6 db and phase margins for unity gain are ± 30 degrees. The ellipse is given by

$$\frac{(1.25 - |\frac{1}{K}| \cos \phi)^2}{(0.75)^2} + \frac{(|\frac{1}{K}| \sin \phi)^2}{(0.5815)^2} = 1 \quad (80)$$

Use of this approximation penalizes gains on the low side of nominal less severely than the unapproximated curve. The approximation is plotted as dotted lines in Figure 13b.

4.2.3.3.4 Complex Gain Margin in Root Loci. The use of a complex gain margin in root locus studies follows immediately. When it is remembered that the minus-one point is a mapping of all system closed-loop roots in the GH-plane, it follows that points in the immediate vicinity of the roots in the s-plane map into points near minus-one in the GH-plane. In fact, if the points corresponding to the closed curve in the complex K-plane (that is, points having the complex coordinates $Ke^{j\phi}$) are plotted on the constant gain and constant phase loci on a root locus plot, the resulting curve surrounding the root in question maps into the curve enclosing the zone of exclusion in the GH-plane. If the region thus defined in the s-plane encloses a portion of the $j\omega$ axis in the vicinity of the root between $j\omega_1$ and $j\omega_2$, the Nyquist curve in the GH-plane lies inside the zone of exclusion for frequencies between ω_1 and ω_2 . Both would be interpreted as violations of the complex gain margin criteria. Figure 14 shows a root locus plot in the vicinity of a parasitic mode root with the complex gain margin of Figure 12a. If this root were more isolated from the surrounding poles and zeros, the phase and gain loci would be very close to straight radial lines and circles, and the complex gain margin on the root locus plot would be identical in shape to the plot on the complex K-plane.

The convenient use of the complex gain margin as an analytical tool rests on two premises. First, that a reasonable basis exists for choosing the form for $Ke^{j\phi}$ (or it is specified), and second, that the engineer can employ it without unduly increasing his computational load. The first question requires consideration of the uncertainties in the parameters having the greatest influence on closed-loop root location. Response requirements, flight control system parameter tolerances, and the uncertainties associated with the mathematical model are the reasons margin is desired in the first place, presumably to avoid complex questions associated with system sensitivity to these variations. The position taken here is that such a basis of choice exists, or that it has been arbitrarily defined by the contracting agency. The second question is more easily answered. Given $Ke^{j\phi}$ in the form of a plot, the zone of exclusion on the Nyquist diagram or on the root locus plot can be constructed graphically. If it takes the form of an ellipse, such as in Figure 13b, the engineer can free-hand sketch an ellipse with error small enough for all but design verification studies requiring relatively high accuracy. The result is a means for avoiding the situation typified by Figure 11 and thus ensure system adequacy.

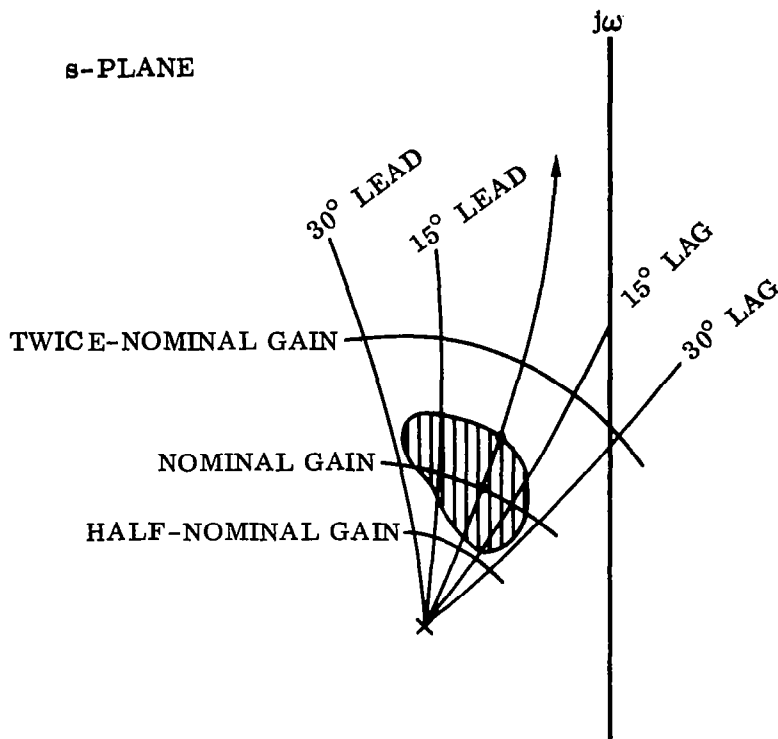


Figure 14. Complex Gain Margin Applied to Root Loci Plots

4.2.3.3.5 Extensions of the Nyquist Technique. Another means utilizing the Nyquist criterion for determination of relative stability consists of changing the form of the closed curve C in the s -plane which is mapped on the GH -plane. One such change would be the mapping of $GH(\sigma + j\omega)$ where σ is so chosen as to exclude roots in the vicinity of the imaginary axis. This has little utility in the study of launch vehicle flight control since the system is dominated by such roots. However, this kind of mapping, if done by machine, provides a means for plotting the orthogonal net discussed in Paragraph 4.2.3.3.1, thus eliminating a source of error. Note that $GH(\sigma + j\omega)$ is no longer the open-loop frequency response, and experimental data for certain elements in the control loop can no longer be used.

This same idea can be used in plotting $GH(-\zeta\omega + j\omega)$; the result is a mapping of a line of constant damping ratio in the GH -plane. The same comments apply.

Another technique given wide coverage in the literature is the M -criterion. A series of circles is drawn in the GH -plane according to

$$M(GH) = \frac{|GH|}{|1 + GH|} \quad (81)$$

These circles surround (but are not centered on) the origin and the minus-one point. Damping of the major roots is inferred from the value of M corresponding to the circle tangent to the Nyquist curve. Unfortunately, the correlation is good for only simple systems and is strictly applicable only to second-order systems having unity feedback. Its utility for application to flight control problems is therefore limited; indeed, if blind use is made of this method of analysis, erroneous answers can result. The reader is referred to the literature for discussions of its limitations and applications (Reference 3).

4.2.3.4 Sensitivity to Parameter Variations. One property a successful flight control system design must have is relative insensitivity in system performance to changes in system parameters. This includes both response and stability properties. The changes arise due to analytical uncertainty (that is, the inability to define the parameters of the mathematical model precisely), statistical variations in the flight environment (due to trajectory perturbations), and components within the flight control system itself.

The first approach to the problem would be to determine the rate of change to a certain property of the system (peak overshoot, dominant root damping ratio, relative stability, open-closed-loop transfer function, etc.) with respect to the particular parameter of interest (rate gain, aerodynamic instability parameter, actuation lag, etc.). If several system parameters of importance are to be considered, one would determine the first variation with respect to the several parameters. However, for most applications, a sensitivity measure of this type suffers from its being an absolute measure when the relative measure is of interest. In short, what is desired is a measure of the relative changes in system properties with respect to a relative change in system parameters; thus

$$S_x^T = \frac{\Delta T}{T} \bigg/ \frac{\Delta x}{x} \quad (82)$$

where the ΔT and Δx are the incremental changes in the property, T , and the parameter, x . In the limit as Δx and ΔT become small,

$$S_x^T = \frac{\partial T/T}{\partial x/x} = \frac{\partial [\ln T]}{\partial [\ln x]} \quad (83)$$

This will be defined as the sensitivity of T with respect to x . As an illustration, consider the block diagram of Figure 3 where the forward loop transfer function has a gain K associated with it; and the feedback loop has a gain k . The overall transmission (the system parameter of interest) is

$$\begin{aligned} T(s) &= \frac{C(s)}{R(s)} \\ &= \frac{KG(s)}{1 + kKGH(s)} \end{aligned} \quad (84)$$

The sensitivity parameters are (from the definition)

$$\left. \begin{aligned} S_K^T &= \frac{1}{1 + kKGH(s)} \\ S_k^T &= \frac{kKGH(s)}{1 + kKGH(s)} \end{aligned} \right\} \quad (85)$$

Thus we have a demonstration of a fundamental property of feedback control systems: that the sensitivity to changes in the forward loop gain approaches zero as the loop gain becomes large and that the sensitivity to changes in the feedback gain approaches a one-to-one correspondence as the loop gain becomes large. This also holds true if we are dealing with frequency dependent compensation within the frequency range where the open-loop system gain is large with respect to unity.

Suppose this same approach is applied to the response of the simplified vehicle-and-autopilot to a gust disturbance (Equation 17). We obtain

$$S_{\mu_\beta}^T = \frac{\mu_\beta}{s^2 + K_A K_R \mu_\xi s + (K_A \mu_\xi - \mu_\beta)} \quad (86)$$

where $T = \frac{\beta}{\beta_W}(s)$. Thus the sensitivity of the gust response to changes in the aerodynamic instability parameter is directly proportional to the latter and inversely proportional to the loop gain as the gain becomes large.

In Paragraph 4.2.3.3 the question was raised of choosing the complex gain margin. There were three considerations mentioned: the response requirements (minimum allowable relative stability for satisfactory response), the analytical uncertainties in the mathematical model which affect the closed-loop root location, and the statistical uncertainties in the mathematical model due to parameter uncertainties, chiefly in the flight control system. The range of closed-loop root locations that result from the latter can be derived based upon known statistical properties of the system random parameters, as follows.

The open-loop transfer function is given by

$$GH(s) = K(x_i, s) e^{j\phi(x_i, s)} \quad i = 1, 2, \dots, n \quad (87)$$

where

x_i = system parameters subject to variation

ϕ = phase angle of GH(s) in radians

K = magnitude (absolute value) of GH(s)

It is desired to determine the relative change in the system gain as a function of the system parameters. From the definition of sensitivity,

$$\begin{aligned}\Delta \ln [GH(s)] &= \sum_{i=1}^n S_{x_i}^{GH} \Delta \ln x_i \\ &= \sum_{i=1}^n \frac{\overline{\partial \ln [GH(s)]}}{\partial \ln x_i} \Delta \ln x_i\end{aligned}\quad (88)$$

where Δ indicates a small incremental change in the pertinent variables and where the bar over the partial derivatives indicates evaluation of the nominal values of the x_i , \bar{x}_i , and the complex frequency \bar{s} corresponding to the system root for which the range of root locations is being obtained. Equation 88 may be developed as

$$\begin{aligned}\Delta \ln [GH(s)] &= \sum_{i=1}^n \frac{\overline{\partial \ln [GH(s)]}}{\partial x_i} \left(\frac{\overline{\partial \ln x_i}}{\partial x_i} \right)^{-1} \Delta \ln x_i \\ &= \sum_{i=1}^n \frac{\overline{\partial \ln [GH(s)]}}{\partial x_i} \bar{x}_i \Delta \ln x_i \\ &= \sum_{i=1}^n g_i \bar{x}_i \Delta \ln x_i + j \sum_{i=1}^n h_i \bar{x}_i \Delta \ln x_i \\ &= u + jv \\ &= \ln K + j\phi\end{aligned}\quad (89)$$

The $\bar{x}_i \Delta \ln x_i$ are considered to be statistically independent Gaussian random variables having zero mean and variance $(\sigma_i)^2$. Thus u and v are likewise Gaussian random variables, being linear functions of $\bar{x}_i \Delta \ln x_i$. They are correlated, since they are both functions of the same random variables and have a bivariate Gaussian probability density function in the uv -plane given by

$$p(u, v) = \frac{e^{-\frac{1}{2} \left[\frac{\sigma_v^2 u^2 - 2\mu_{uv} uv + \sigma_u^2 v^2}{\sigma_u^2 \sigma_v^2 - \mu_{uv}^2} \right]}}{2\pi (\sigma_u^2 \sigma_v^2 - \mu_{uv}^2)^{1/2}} \quad (90)$$

The quantities u and v represent the incremental changes in, respectively, the natural logarithm of the gain and the phase of the open-loop transfer function $GH(s)$ evaluated at \bar{s} and \bar{x}_i , $i = 1, 2, \dots, n$, due to incremental changes in the system parameters x_i . Since the variance of a linear combination of statistically independent random variables is equal to the sum of the same linear combination of the variances, one has

$$\left. \begin{aligned} \sigma_u^2 &= \sum_{i=1}^n g_i^2 \sigma_i^2 \\ \sigma_v^2 &= \sum_{i=1}^n h_i^2 \sigma_i^2 \\ \mu_{uv} &= \sum_{i=1}^n g_i h_i \sigma_i^2 \end{aligned} \right\} \quad (91)$$

where the latter quantity is the covariance of u and v ; that is, the expected value of the product of u and v . Since u and v are correlated, this quantity is not generally zero.

The integral of $p(u, v)$ over a region R of the uv -plane represents the joint probability that both u and v are within R . For purposes of defining the range of closed-loop root locations, a region R is sought for which the probability of u and v being within that region is 99.73 percent; that is, the three-sigma condition. There is an infinite number of such regions, but it seems most logical to pick the one for which the probability density on the boundary of the region is constant. Thus the probability of u and v being just inside R is independent of the point on the boundary chosen. Referring to Equation 90, the exponent of e is constant when u and v satisfy an equation corresponding to an ellipse. The presence of the uv term indicates that the ellipse has been rotated with respect to the normal orientation where the equation would contain no such term.

Suppose a new coordinate system having axes w and z is defined to be rotated by a counterclockwise angle, γ , with respect to the uv axes system. The coordinate transformation is:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos \gamma & -\sin \gamma \\ \sin \gamma & \cos \gamma \end{bmatrix} \begin{bmatrix} w \\ z \end{bmatrix} \quad (92)$$

The variances of the old variables are given in terms of the new by

$$\left. \begin{aligned} \sigma_u^2 &= \sigma_w^2 \cos^2 \gamma + \sigma_z^2 \sin^2 \gamma \\ \sigma_v^2 &= \sigma_w^2 \sin^2 \gamma + \sigma_z^2 \cos^2 \gamma \\ \mu_{uv} &= (\sigma_w^2 - \sigma_z^2) \cos \gamma \sin \gamma + \mu_{wz} (\cos^2 \gamma - \sin^2 \gamma) \end{aligned} \right\} \quad (93)$$

In the new coordinate system, w and z are to appear uncorrelated, whence μ_{wz} vanishes. The resulting system of equations can then be solved for γ , σ_w^2 , and σ_z^2 , giving

$$\left. \begin{aligned} \tan 2\gamma &= 2\mu_{uv}/(\sigma_u^2 - \sigma_v^2) \\ \sigma_w^2 &= (\sigma_u^2 - \sigma_v^2 \tan^2 \gamma)/(1 - \tan^2 \gamma) \\ \sigma_z^2 &= (\sigma_v^2 - \sigma_u^2 \tan^2 \gamma)/(1 - \tan^2 \gamma) \end{aligned} \right\} \quad (94)$$

The probability density function for w and z is given by

$$p(w, z) = \frac{e^{-\frac{1}{2} \left[\frac{w^2}{\sigma_w^2} + \frac{z^2}{\sigma_z^2} \right]}}{2\pi\sigma_w\sigma_z} \quad (95)$$

The ellipse surrounding the region R for which the probability of w and z being within R is 99.73 percent and satisfying the criterion of constant probability density on the boundary is given by

$$\begin{aligned} \frac{w^2}{(3.44\sigma_w)^2} + \frac{z^2}{(3.44\sigma_z)^2} &= \frac{w^2}{a^2} + \frac{z^2}{b^2} \\ &= 1 \end{aligned} \quad (96)$$

where the value 3.44 arises from integrating $p(w, z)$ over the elliptical region represented by the exponent of e in Equation 95 equaling some constant, the constant to be

determined such as to give the required probability. When transformed back into uv coordinates the ellipse when plotted as phase ($\phi = 57.3$ v) in degrees versus gain ($\delta = 8.68$ u) in decibels becomes

$$\left(\frac{\delta}{8.68}\right)^2 \left[\frac{\cos^2 \gamma}{a^2} + \frac{\sin^2 \gamma}{b^2} \right] + 2 \left(\frac{\delta}{8.68}\right) \left(\frac{\phi}{57.3}\right) \cos \gamma \sin \gamma \left[\frac{1}{a} - \frac{1}{b} \right] + \left(\frac{\phi}{57.3}\right)^2 \left[\frac{\sin^2 \gamma}{a^2} + \frac{\cos^2 \gamma}{b^2} \right] = 1 \quad (97)$$

When this ellipse is transformed back into the complex K-plane, the required range of expected root locations is determined. This can then be applied as a criterion on root locus plots or on the Nyquist diagram. Note that the curve obtained in the complex K-plane is identical to the zone of exclusion in the Nyquist plot; the curve is rotated about the origin by 180 degrees. This property results from the manner in which the curve was derived. To recapitulate, the procedure is as follows:

1. Define the variances $(\sigma_i)^2$ of the system parameters. This is really the heart of the problem and requires exercise of engineering judgment. If flight control system parameter variations are being considered one can assume the parameters to be normally distributed about the nominal, in which case the tolerances correspond to three-sigma figures; or one can take a more jaundiced view and assume the probability distribution to be constant over the tolerance range, in which case the tolerance divided by $\sqrt{3}$ becomes the σ_i . This assumption is based on the Central Limit Theorem which states that u and v will tend to be normally distributed if a large number of non-normally distributed parameters are considered. The σ_i are the square roots of the variances of the non-normal probability density functions for the x_i .
2. Determine the g_i and h_i for each parameter. This is not too difficult, for generally speaking, $GH(s)$ is a ratio of products of a number of factors, in which case $\ln GH(s)$ becomes a summation of logarithms and phase angles, each dependent on one of the parameters. The g_i and h_i can then be written by inspection and then evaluated at \bar{x}_i and \bar{s} . The complex frequency used is that of the nominal closed-loop root location, and can be inferred from the Nyquist diagram or root locus plot. This can also be done graphically, using root locus techniques. First, plot constant gain and constant phase loci for a range of gain and phase that will encompass the region of possible root locations; then determine the root locations corresponding to a variation of σ_i in each of the x_i about nominal. Each root location corresponds to a particular $g_i \sigma_i$ and $h_i \sigma_i$ as read off the orthogonal net of the constant gain and phase loci.

3. Evaluate σ_u , σ_v , and μ_{uv} .
4. Evaluate γ , σ_w , σ_z , a, and b.
5. Evaluate $Ke^{j\phi}$, the complex gain margin. This can be done in several ways. One can evaluate the coefficients of Equation 97, solve for δ (equal to $20 \log_{10} K$), and plot directly on the complex K-plane. Or the ellipse of Equation 96 can be plotted (using equal scales for w and z), overlaid with a semitransparent piece of paper rotated clockwise through the angle γ , and traced: the result is a plot of u in nepers (horizontal scale) versus phase in radians (vertical scale). Rescaling to decibels (multiply the u coordinate values by 8.68) and degrees (multiply the v coordinates by 57.3), the result can be mapped to the complex K-plane. And, of course, one can take the easy way out and program the whole procedure on a digital computer.

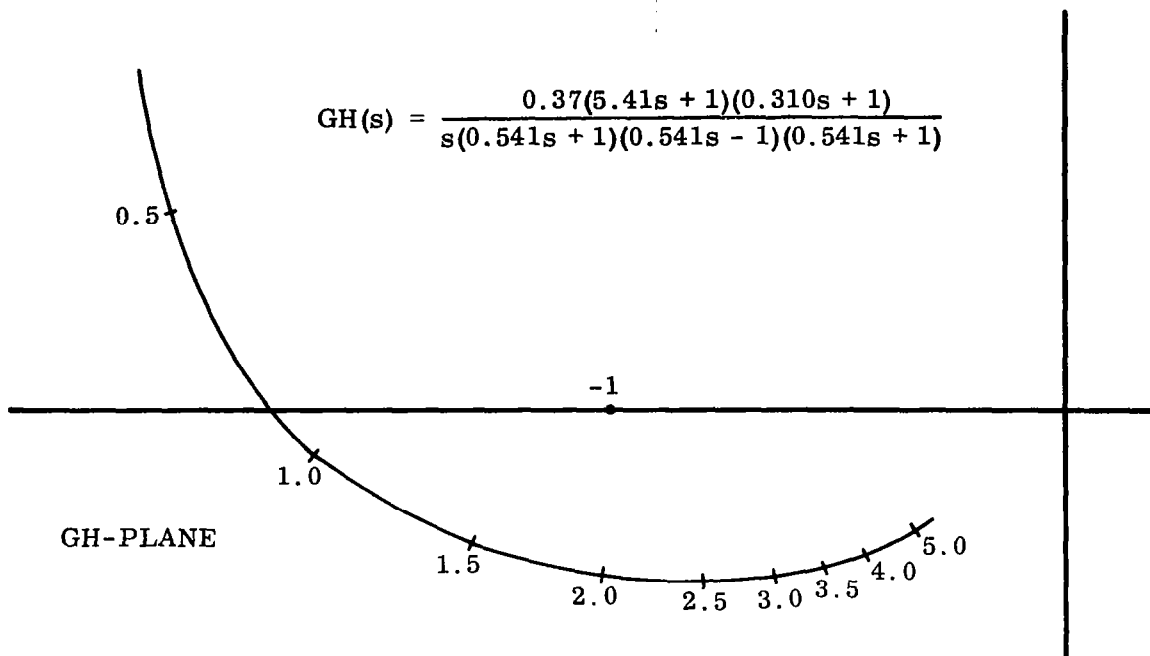
The results of calculating g_i and h_i should not be overlooked. The relative magnitudes of each term in the summation making up the variances σ_u^2 and σ_v^2 give one an idea of the relative importance each parameter variation has on the overall system root being considered. One may find that only one parameter has any importance, in which case the ellipse in the uv-plane becomes a straight line. When mapped on the K-plane, the result is a curved line; when further mapped on the root locus plot or Nyquist diagram, one has an approximation to the root motion with respect to the particular parameter involved. It would be exact except for the fact that Equation 89 is a linear approximation (\bar{s} is a fixed quantity).

4.2.3.5 Example - Rigid Vehicle Flight Control. The simplified system of Paragraph 4.2.1 has the open-loop transfer function,

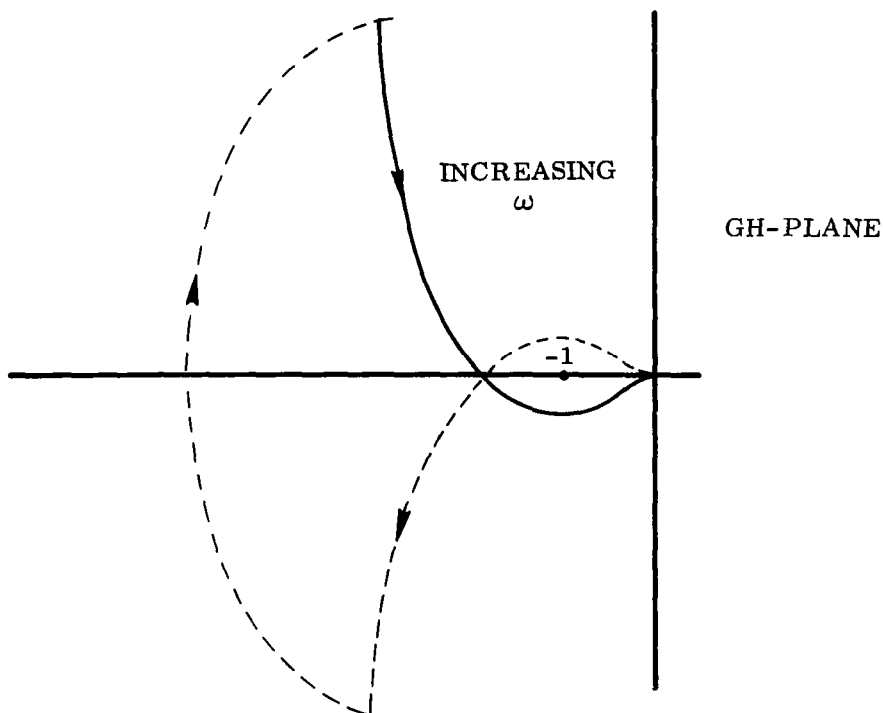
$$GH(s) = \frac{K_A \mu_\xi K_I \left(\frac{s}{K_I} + 1 \right) (K_R s + 1)}{\mu_\beta s \left(\frac{s^2}{\mu_\beta} - 1 \right) (\tau_C s + 1)} \quad (98)$$

For the system parameters of Tables 1 and 2, the Nyquist plot of the open-loop frequency response is given in Figure 15. The number of clockwise encirclements of the minus-one point is -1; and since there is one pole of the open-loop transfer function in the right-half s-plane the system is seen to be stable.

The phase margin is measured as 22.5 degrees, the negative gain margin as 0.58 or 4.7 db. In more precise modeling of the vehicle dynamics there would be a positive gain margin as well, but in this simplified system the lag at high frequencies approaches -180 degrees for positive ω , resulting in stable roots regardless of the gain. The diagram also shows the rigid body root going unstable at $\omega \cong 0.9$ rad/sec, which is in agreement with the root locus of Figure 6.



a. Nyquist Plot for Positive ω



b. Showing Number of Encirclements

Figure 15. Nyquist Diagram of Rigid Vehicle Flight Control System

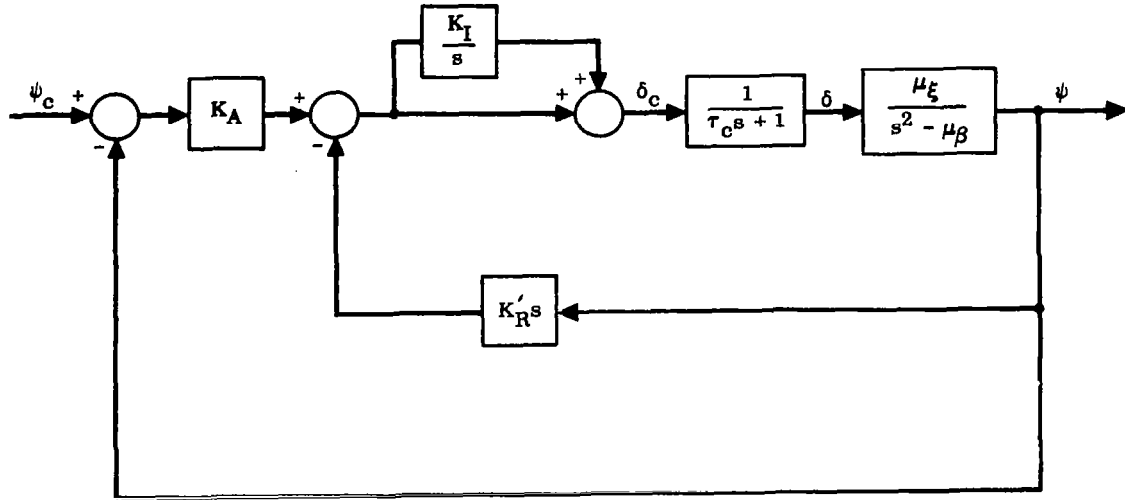
This figure is also an example of where it is difficult to estimate the dominant root location with any degree of accuracy. The frequency gradient along the curve is changing fairly rapidly and the minus-one point is well away from the curve.

The range of root locations required for system statistical parameter variations is developed as follows. The variable parameters with their associated tolerances or ranges of variation are given in Table 3. The autopilot gains are correlated as given here. The physical reason for this is that the autopilot configuration is in reality that given in Figure 16. The end-to-end rate gain is formed with one summing resistor, the end-to-end position gain with another. The percentage tolerance on the integral gain is now given by

$$\begin{aligned}
 \frac{\sigma_I^2}{K_I^2} &= \frac{\sigma_{K_A K_I}^2}{(K_A K_I)^2} - \frac{\sigma_{K_A}^2}{K_A^2} \\
 &= (0.25)^2 - (0.15)^2 \\
 &= (0.20)^2
 \end{aligned} \tag{99}$$

Table 3. Ranges of Parameter Variation for Example Case - Rigid Vehicle Flight Control

PARAMETER	TOLERANCE	NOMINAL VALUE
K_A	$\pm 15\%$	1.75
$K_A K_R = K_R'$	$\pm 12\%$	0.543 ($K_R = 0.310$)
$K_A K_I$	$\pm 25\%$	0.324 ($K_I = 0.185$)
τ_c	$\pm 60\%$	0.054
μ_ξ	$\pm 6\%$	3.901
μ_β	$\pm 10\%$	3.407



$$GH(s) = \frac{\mu_{\xi}(K'_R s + K_A)(s + K_I)}{s(s^2 - \mu_{\alpha})(\tau_c s + 1)}$$

Figure 16. Autopilot Block Diagram Showing Physical Placement of Parameters

where the variances have been assumed in the same proportional relationship to the tolerances for all three variables.

Taking the logarithm of $GH(s)$,

$$\begin{aligned} \ln[GH(s)] = & \ln \mu_{\xi} + \frac{1}{2} \ln \left[(K_I + \sigma)^2 + \omega^2 \right] + \frac{1}{2} \ln \left[(K'_R \sigma + K_A)^2 + K'^2_R \omega^2 \right] \\ & - \frac{1}{2} \ln \left[\sigma^2 + \omega^2 \right] - \frac{1}{2} \ln \left[(\sigma^2 - \omega^2 - \mu_{\beta})^2 + 4\sigma^2 \omega^2 \right] \\ & - \frac{1}{2} \ln \left[(\tau_c \sigma + 1)^2 + \tau_c^2 \omega^2 \right] + j \left[\tan^{-1} \left(\frac{\omega}{\sigma + K_I} \right) + \tan^{-1} \left(\frac{K'_R \omega}{K'_R \sigma + K_A} \right) \right. \\ & \left. - \tan^{-1} \left(\frac{\omega}{\sigma} \right) - \tan^{-1} \left(\frac{2\sigma \omega}{\sigma^2 - \omega^2 - \mu_{\beta}} \right) - \tan^{-1} \left(\frac{\tau_c \omega}{\tau_c \sigma + 1} \right) \right] \end{aligned} \quad (100)$$

The g_i and h_i are given by

$$\left. \begin{aligned}
 g_{u_\xi} &= \frac{1}{\bar{\mu}_\xi} & h_{\mu_\xi} &= 0 \\
 g_{K_I} &= \frac{\bar{K}_I + \bar{\sigma}}{(\bar{K}_I + \bar{\sigma})^2 + \bar{\omega}^2} & h_{K_I} &= -\frac{\bar{\omega}}{(\bar{K}_I + \bar{\sigma})^2 + \bar{\omega}^2} \\
 g_{K'_R} &= \frac{\bar{K}'_R(\bar{\sigma}^2 + \bar{\omega}^2) + \bar{K}_A \bar{\sigma}}{(\bar{K}'_R \bar{\sigma} + \bar{K}_A)^2 + \bar{K}'_R{}^2 \bar{\omega}^2} & h_{K'_R} &= \frac{\bar{K}_A \bar{\omega}}{(\bar{K}'_R \bar{\sigma} + \bar{K}_A)^2 + \bar{K}'_R{}^2 \bar{\omega}^2} \\
 g_{K_A} &= \frac{\bar{K}'_R \bar{\sigma} + \bar{K}_A}{(\bar{K}'_R \bar{\sigma} + \bar{K}_A)^2 + \bar{K}'_R{}^2 \bar{\omega}^2} & h_{K_A} &= -\frac{\bar{K}'_R \bar{\omega}}{(\bar{K}'_R \bar{\sigma} + \bar{K}_A)^2 + \bar{K}'_R{}^2 \bar{\omega}^2} \\
 g_{\tau_c} &= -\frac{\bar{\tau}_c(\bar{\sigma}^2 + \bar{\omega}^2) + \bar{\sigma}}{(\bar{\tau}_c \bar{\sigma} + 1)^2 + \bar{\tau}_c{}^2 \bar{\omega}^2} & h_{\tau_c} &= -\frac{\bar{\omega}}{(\bar{\tau}_c \bar{\sigma} + 1)^2 + \bar{\tau}_c{}^2 \bar{\omega}^2} \\
 g_{\mu_\beta} &= \frac{(\bar{\sigma}^2 - \bar{\omega}^2 - \bar{\mu}_\beta)}{(\bar{\sigma}^2 - \bar{\omega}^2 - \bar{\mu}_\beta)^2 + 4\bar{\sigma}^2 \bar{\omega}^2} & h_{\mu_\beta} &= \frac{1}{(\bar{\sigma}^2 - \bar{\omega}^2 - \bar{\mu}_\beta)^2 + 4\bar{\sigma}^2 \bar{\omega}^2}
 \end{aligned} \right\} \quad (101)$$

Assuming the probability density functions to be normal distributions with the three-sigma deviation corresponding to the extremes of the tolerance band, the σ_i are given by one-third of the tolerance. Evaluating g_i , h_i , and σ_i for $\bar{s} = \bar{\sigma} + j\bar{\omega} = -0.754 + j1.732$, the rigid body mode closed-loop root, the values of Table 4 are obtained.

From these values the following result for the parameters of the zone of exclusion.

$$\left. \begin{aligned}
 \sigma_u^2 &= 0.002648 & \sigma_u &= 0.05147 \\
 \sigma_v^2 &= 0.001997 & \sigma_v &= 0.04465 \\
 \mu_{uv} &= -0.001382 \\
 \gamma &= -38.38 \text{ deg} \\
 \sigma_w^2 &= 0.00375 & \sigma_w &= 0.0612 \\
 \sigma_z^2 &= 0.000901 & \sigma_z &= 0.03005
 \end{aligned} \right\} \quad (102)$$

Table 4. Evaluation of g_1 , h_1 , and σ_1 - Example Case

PARAMETER	g_1	h_1	σ_1
$\bar{K}_A = 1.75$	0.4995	-0.3505	0.0875
$\bar{K}'_R = 0.543 \text{ sec}$	0.2302	1.130	0.0217
$\bar{K}_I = 0.185 \text{ sec}^{-1}$	-0.1710	-0.521	0.01234
$\bar{\tau}_c = 0.054 \text{ sec}$	0.605	-1.866	0.1080
$\bar{\mu}_\xi = 3.901 \text{ sec}^{-2}$	0.2562	0	0.07805
$\bar{\mu}_\beta = 3.407 \text{ sec}^{-2}$	-0.1429	0.02447	0.1136

In the course of evaluating the contributors to α_w and α_v it is evident that the position gain tolerance contributes most heavily to the overall variation in root location. The rate gain and actuation lag are relatively less important, while the effect of the integrator tolerance and dispersions in inertial and aerodynamic properties contribute almost nothing. The zone of exclusion is defined by (δ in decibels, ϕ in degrees)

$$0.6645\delta^2 + 0.1394\delta\phi + 0.0202\phi^2 = 1 \quad (103)$$

This region has a maximum gain deviation of ± 1.54 db, a phase deviation of ± 8.83 deg, and shows strong correlation with increasing gain going with increasing lag and vice versa. When used as a criterion for system stability it is seen that the system has substantial margin stability-wise for the parameter range considered. However, the response properties are equally important. Under the extremes of this range of root locations the damping could decrease to 0.25, which is probably insufficient for well-damped response to atmospheric disturbances.

This difficulty can be alleviated by an increase in the rate gain. For computational purposes, one calculates the gain and phase of

$$G_c(s) = \frac{0.45s + 1}{0.31s + 1} \quad (104)$$

which when combined with the data already given for the Nyquist diagram for $K'_R = 0.31$ sec, yields the open-loop frequency response for $K'_R = 0.45$ sec. This plot is the lower of the two in Figure 17. Strictly speaking, the range of root locations should also be recomputed for the new K'_R , but the change is small enough to have relatively little effect on its dimensions. One disadvantage of the higher rate gain is the increased

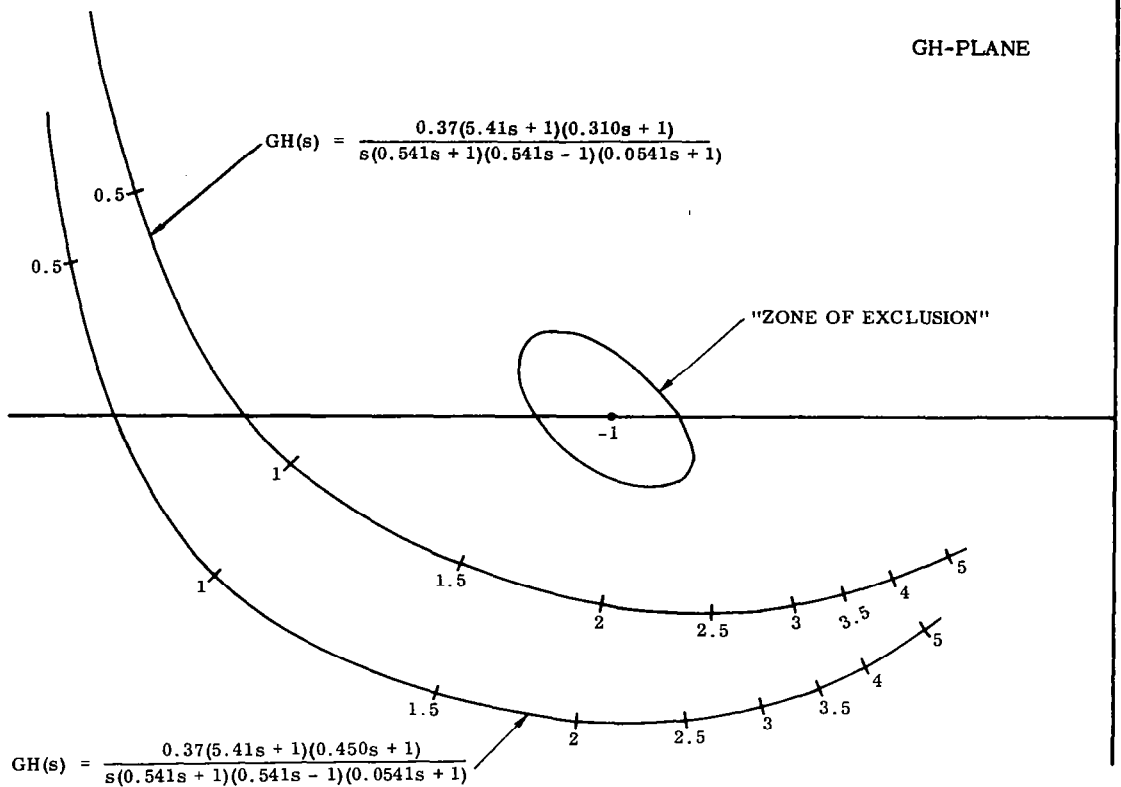


Figure 17. Nyquist Diagram of Example System with Zone of Exclusion

gain at high frequencies. While this is of no consequence for this simplified model, it could cause trouble when the higher modes of response are considered.

4.2.4 AUXILIARY FEEDBACK LOOPS FOR RIGID BODY CONTROL. In the flight control of large space launch vehicles there is a strong incentive to minimize the flight loads during the ascent through the atmosphere. If this can be done successfully, the weight of the structure can accordingly be reduced to maximize the vehicle performance. There are several schemes that might be employed.

First, one can pre-program the flight control system commands in such a fashion that the boost phase terminal objectives (altitude, flight path angle, and velocity) are met while at the same time the inflight loads are reduced. This requires prior knowledge of the atmospheric environment (wind velocity and direction as a function of altitude). To implement this scheme, one would release a balloon prior to launch, measure the wind profile from radar tracking data, and use the information thus obtained to derive the optimal sequencing of booster phase commands (this requires a digital computer routine incorporating terminal objectives and loading constraints).

Second, one could use an onboard inertial guidance system in implementing a set of guidance equations designed to minimize booster phase flight loads. The vehicle

accelerations induced by winds aloft are sensed by the guidance system and used to guide the vehicle in such a fashion that the boost phase terminal objectives are met while reducing the lateral loads. This is fundamentally a feedback of the lateral acceleration.

Third, auxiliary body-fixed sensors (either angle-of-attack instrumentation or accelerometers sensitive to lateral acceleration) are used as sources of the loading information. The signals generated are used to modify the attitude commands such as to reduce the inflight loads.

Each of these schemes or a combination of them implies some compromise between the desire for reducing the inflight loads and meeting the boost phase terminal objectives with sufficient accuracy such that performance is not degraded. The degree of tradeoff is not easily determined; usually it involves detailed trajectory studies. However, it is safe to say that some compromise can be worked out that will show substantial improvement over what can be obtained with no efforts made in this direction. The price paid is in system complexity and consequent degradation of flight control system reliability. The discussion that follows assumes the third alternative above; the first and second are subjects for performance, trajectory, and guidance systems optimization. The use of auxiliary sensors and feedback signals modifies the vehicle flight control system stability qualities, thus falling within the subject matter of this monograph.

Consider Equation 1 where the velocity is assumed high enough to result in the velocity-dependent terms going to zero. The thrust vector deflection angle is given by the control law,

$$\xi(s) = G_c(s)\psi_c(s) - G_\psi(s)\psi(s) - G_y(s)\ddot{y}(s) - G_\beta(s)\beta(s) \quad (105)$$

where G represents frequency-dependent transfer functions relating to commanded attitude, actual attitude, lateral acceleration, and sideslip angle. If β is taken as a measure of the aerodynamic load, one has the response,

$$\beta(s) = \frac{G_c\mu_\xi\psi_c(s) + \left[\left(1 + \frac{T_c}{M}G_y\right)s^2 + G_\psi\mu_\xi\right]\beta_W(s)}{\left(1 + \frac{T_c}{M}G_y\right)s^2 - \mu_\beta + G_\psi\mu_\xi + G_y\frac{F_\beta}{M}\left(1 + \frac{l_\beta}{l_c}\right)\mu_\xi + G_\beta\mu_\xi} \quad (106)$$

where the frequency dependence of G is understood. It is clear that proper tailoring of $\psi_c(t)$ will result in a minimum $\beta(t)$ if $\beta_W(t)$ is known (the first alternative mentioned above).

4.2.4.1 Angle-of-Attack Feedback. Suppose an angle-of-attack sensor is used. In this case,

$$\left. \begin{aligned} G_{\psi}(s) &= \frac{K_A (s + K_I)(K_R s + 1)}{s (\tau_c s + 1)} \\ G_{\beta}(s) &= \frac{K_A K_{\beta} (s + K_I)}{s (\tau_c s + 1)} \end{aligned} \right\} \quad (107)$$

where K_{β} represents the ratio of the gain associated with the auxiliary loop to the position gain. The response becomes ($\psi_c(t) = 0$)

$$\begin{aligned} \frac{\beta}{\beta_W}(s) &= \frac{K_A \mu_{\xi} K_R s^2 + K_A \mu_{\xi} (1 + K_R K_I) s + K_A \mu_{\xi} K_I}{\tau_c s^4 + s^3 + (K_A \mu_{\xi} K_R - \mu_{\beta} \tau_c) s^2 + [K_A \mu_{\xi} (1 + K_{\beta} + K_R K_I) - \mu_{\beta}] s} \\ &\quad + K_A \mu_{\xi} [1 + K_{\beta}] K_I \end{aligned} \quad (108)$$

It is seen that the steady-state response to an input step is reduced by $(1 + K_{\beta})^{-1}$. On comparing the characteristic equation here with that of the conventional autopilot (Equation 5) it is noted the coefficients are the same when:

<u>Conventional Autopilot</u>	=	<u>Autopilot Employing Angle-of-Attack Feedback</u>
K_A	=	$K_A (1 + K_{\beta})$
K_R	=	$\frac{K_R}{1 + K_{\beta}}$

This implies that the addition of angle-of-attack feedback requires a reduction of the proportional gain K_A to result in the same degree of stability. The end-to-end rate gain remains the same. The response to guidance commands will be characterized by steady-state position error, unlike the conventional autopilot, because of the auxiliary loop.

4.2.4.2 Accelerometer Feedback. Another choice of auxiliary sensor is an accelerometer sensitive to the lateral (perpendicular to the body axes) acceleration. The accelerometer is mounted a distance ℓ_a forward of the vehicle center of gravity, resulting in a sensed acceleration,

$$\ddot{y}_a = \ddot{y} + \ell_a \ddot{\psi} \quad (109)$$

The frequency dependent transfer functions are

$$\left. \begin{aligned} G_{\psi}(s) &= \frac{K_A (s + K_I) (K_y \ell_a s^2 + K_R s + 1)}{s (\tau_c s + 1)} \\ G_y(s) &= K_y \frac{K_A (s + K_I)}{s (\tau_c s + 1)} \end{aligned} \right\} \quad (110)$$

where K_y is the ratio of the auxiliary loop gain to the position gain. The characteristic equation becomes

$$\begin{aligned} &\tau_c s^4 + \left[1 + K_A \mu_{\xi} K_y (\ell_a - \ell_{cp}) \right] s^3 + \left\{ K_A \mu_{\xi} \left[K_R + K_I K_y (\ell_a - \ell_{cp}) \right] - \mu_{\beta} \tau_c \right\} s^2 \\ &+ \left\{ K_A \mu_{\xi} \left[1 + K_y \frac{F\beta}{M} \left(1 + \frac{\ell_{\beta}}{\ell_c} \right) + K_R K_I \right] - \mu_{\beta} \right\} s \\ &+ K_A \mu_{\xi} \left[1 + K_y \frac{F\beta}{M} \left(1 + \frac{\ell_{\beta}}{\ell_c} \right) \right] K_I = 0 \end{aligned} \quad (111)$$

where $\ell_{cp} = \frac{I}{M \ell_c}$, the distance between the center of gravity and the center of percussion (the instantaneous center of rotation). This equation is quite formidable to handle using the Routh-Hurwitz criterion. However, we note that:

1. An accelerometer location aft of the center of percussion is destabilizing, causing positive feedback.
2. The stability qualities of the system of low frequencies are identical to those of angle-of-attack feedback if:
 - a. The accelerometer is located at the center of percussion.
 - b. The accelerometer loop gain is set equal to the angle-of-attack loop gain divided by the quantity $\frac{F\beta}{M} \left(1 + \frac{\ell_{\beta}}{\ell_c} \right)$.
3. The steady-state response of the angle of attack to a step input $\beta_W(t)$ is reduced by a factor $\left[1 + K_y \frac{F\beta}{M} \left(1 + \frac{\ell_{\beta}}{\ell_c} \right) \right]^{-1}$.

The same remarks as before apply: use of accelerometer feedback implies a reduction of the end-to-end position gain to result in the same stability qualities.

4.2.4.3 Minimum Drift Gains. One popular compromise between the load-relieving qualities of auxiliary feedback loops and the trajectory dispersions suffered is to use minimum drift gains. Referring to the second equation in Figure 1, the lateral acceleration relative to the flight path is minimized when

$$\psi = - \frac{\ddot{y}}{A_x} \quad (112)$$

This requires acceleration feedback with a gain inversely proportional to axial acceleration and can be achieved by using a free pendulum type of accelerometer. The trim angle of the pendulum relative to the vehicle gain gives a signal proportional to the quantity desired. The gain used is thus $K_y = (A_x)^{-1}$, and will result in minimum lateral

drift and a reduction in the steady-state angle of attack by $\left[1 + \frac{F_\beta}{T_c + T_f - D} \left(1 + \frac{l_\beta}{l_c} \right) \right]^{-1}$. Feedback of this type also minimizes trajectory dispersions due to center-of-gravity offsets, thrust misalignments, and the like.

4.2.4.4 Analysis Using D-Decomposition. The use of auxiliary loops to relieve aerodynamic loads increases the dimensionality of the control stability analysis problem. Using the techniques discussed earlier, this generally requires multiple applications, one for each loop. With several trials, a suitable result can be derived. Another approach to this problem is afforded by the technique known as D-decomposition, a popular approach for linear control systems analysis in Russian literature. A description of the technique is given here (following that of Reference 6) because it seems well adapted to the analysis of the multiple-loop flight control system. Unfortunately, it is not as adaptable to hand calculations, even for the simplified rigid body flight control systems under discussion.

Basically, the technique enables definition of the permissible values over which each of several system parameters can vary while maintaining the system roots in a specified portion of the s-plane, usually the entire left-half plane. If each system parameter is considered as a dimension of an N-dimensional parameter space (N equals the number of parameters being considered), the space is decomposed into various regions, each characterized by a particular number of unstable roots in the characteristic equation. Each region is denoted by D (k, n-k), where k is the number of roots lying in the left-half s-plane (or other portion; for example, that portion to the left of $s = \zeta \omega \pm j\omega$) and n is the degree of the characteristic equation. Thus D(n, 0) denotes the region wherein all the system roots have the desired degree of stability, and the boundaries of this region are stability boundaries for the n parameters being considered. The resultant decomposition of the parameter space into regions is called D-Decomposition.

The technique is best understood as an extension or modification of the Nyquist criterion. Consider the open-loop transfer function $KGH(s)$, where K represents the loop gain or the single parameter under discussion. The characteristic equation is then represented by

$$1 + KGH(s) = 1 + K \frac{N(s)}{D(s)} = 0 \quad (113)$$

where $N(s)$ and $D(s)$ are the numerator and denominator polynomials in s , respectively, of the open-loop transfer function. For a particular choice of K , the number of clockwise encirclements of the minus-one point in the KGH -plane determines the stability of the system, given N_p , the number of roots of $D(s) = 0$, lying in the right-half s -plane. This expression can be recast into the form,

$$\frac{D(s)}{N(s)} + K = 0 \quad (114)$$

If the frequency response of $\frac{D(s)}{N(s)}$ is plotted in the $\frac{D(s)}{N(s)} = [GH(s)]^{-1}$ plane the number of clockwise encirclements of point $-K$ determines the stability of the system, given the number of roots of $N(s) = 0$ lying in the right-half s -plane. Since there are several points, in general several segments of the real axis of the $[GH(s)]^{-1}$ plane, for which the criterion may be satisfied the permissible bounds on K are defined for which the system has stable roots. Thus the real axis is decomposed into sets of points, $D(k, n-k)$, of which the set $D(n, 0)$ represents the stable region. The boundaries of these regions correspond to intersections of the Nyquist curves with the real axis.

These intersections are of two types: single intersections (where ω equals 0 or $\pm\infty$) and double (where ω has particular intermediate values). Each segment of the real axis can be tested via the Nyquist criterion for the number of unstable roots, $n - k$. This number increases or decreases by one as K is varied such that it crosses a single intersection, and increases or decreases by two when varied such that it crosses a double intersection. This is termed the index property.

Now consider a characteristic equation of the form,

$$D(s) + \sum_{i=1}^n K_i N_i(s) = 0 \quad (115)$$

This is rewritten as

$$\frac{D(s)}{N_\ell(s)} + \sum_{\substack{i=1 \\ i \neq \ell}}^n \frac{K_i N_i(s)}{N_\ell(s)} + K_\ell = 0 \quad (116)$$

For any choice of $K_i (i \neq \ell)$, the set $D_\ell(n, 0)$ can be defined (if it exists) for the parameter K_ℓ . If the K_i are varied over all possible values, one will define the bounds on the set $D(n, 0)$ in the N -dimensional parameter space of the $K_i (i = 1, 2, \dots, N)$ for which the system is stable.

Stated in this fashion, the technique is not very useful. It enables one to solve for the stability boundaries for the one-dimensional case, but the N -dimensional case is, to say the least, tedious. Fortunately, a relatively simple solution can be obtained for the two-parameter case. Consider the characteristic equation,

$$D(s) + K_1 N_1(s) + K_2 N_2(s) = 0 \quad (117)$$

Setting $s = j\omega$, one can solve for those values of K_1 and K_2 corresponding to intersections of the Nyquist curve with the real axis by equating the real and imaginary parts of Equation 117 to zero. Setting

$$\left. \begin{aligned} D(j\omega) &= D_R(\omega) + j D_I(\omega) \\ N_1(j\omega) &= N_{R_1}(\omega) + j N_{I_1}(\omega) \\ N_2(j\omega) &= N_{R_2}(\omega) + j N_{I_2}(\omega) \end{aligned} \right\} \quad (118)$$

one has

$$\left. \begin{aligned} K_1 &= \frac{D_I N_{R_2} - D_R N_{I_2}}{N_{R_1} N_{I_2} - N_{R_2} N_{I_1}} \\ K_2 &= \frac{D_R N_{I_1} - D_I N_{R_1}}{N_{R_1} N_{I_2} - N_{R_2} N_{I_1}} \end{aligned} \right\} \quad (119)$$

The parameters K_1 and K_2 are thus functions of ω and when plotted on the K_1, K_2 plane they define those boundaries corresponding to double intersections of the Nyquist plot with the real axis. To find single intersections, one sets s equal to zero and/or $\pm\infty$ in Equation 117, which will result in additional boundaries in the K_1, K_2 plane. The regions thus defined are tested by picking particular values for K_1 and K_2 lying within the region and evaluating for stability by any convenient means. The result is the D -decomposition of K_1, K_2 space into regions having various numbers of unstable roots. The relative system stability of a particular operating point K_1, K_2 may be inferred from the distance that this point lies from the boundaries of the region. The relationship is not direct as in Nyquist diagrams because the mapping defined is not conformal.

4.2.4.5 Example - D-Decomposition Analysis of Accelerometer Feedback. The characteristic equation of Equation 111 may be rewritten as

$$s \left[(s^2 - \mu_\beta)(\tau_c s + 1) + K_A \mu_\xi K_R (s + K_I) \right] + K_\psi K_A \mu_\xi (s + K_I) + K_y K_A \mu_\xi (s + K_I) \left[(\ell_a - \ell_{cp}) s^2 + \frac{F\beta}{M} \left(1 + \frac{\ell_\beta}{\ell_c} \right) \right] = 0 \quad (120)$$

where K_ψ represents the gain applied to the position feedback signal ($K_\psi = 1$) in the conventional autopilot). The first term in Equation 120 is recognized as the characteristic equation of a system employing rate feedback only.

For single intersections, K_ψ and K_y both approach infinity as s approaches infinity. For $s = 0$,

$$K_\psi = -K_y \frac{F\beta}{M} \left(1 + \frac{\ell_\beta}{\ell_c} \right) \quad (121)$$

For double intersections the denominators of Equations 119 go to zero. This means that the boundaries of the regions sought are not functions of ω and that a specific relation between K_ψ and K_y can be found. In short, either the real or imaginary parts of the characteristic equation go to zero ($s = \pm j\omega$) independently of K_ψ and K_y . Using this approach we find the imaginary part to be

$$\text{Im(char. eq.)} = j\omega \left[K_R - \frac{(\omega^2 + \mu_\beta)(K_I + \omega^2 \tau)}{K_A \mu_\xi (K_I + \omega^2)} \right] \quad (122)$$

which, when the variables of Tables 1 and 2 are used, yields

$$\omega^2 = 0.2046, 49.852$$

The real part of the characteristic equation is

$$\begin{aligned} \text{Re(char. eq.)} = & \frac{\omega^2 (\tau K_I - 1)(\omega^2 + \mu_\beta)}{K_A \mu_\xi (K_I + \omega^2)} + K_\psi \\ & + K_y \left[\frac{F\beta}{M} \left(1 + \frac{\ell_\beta}{\ell_c} \right) - (\ell_a - \ell_{cp}) \omega^2 \right] \end{aligned} \quad (123)$$

which, when evaluated at the above values for ω^2 using $(\ell_a - \ell_{cp}) = 38.22$ ft, yields

$$\left. \begin{aligned} K_\psi &= 0.4487 - 94.74 K_y \\ &= 7.718 + 1798 K_y \end{aligned} \right\} \quad (124)$$

The decomposition of the K_ψ, K_y plane is given in Figure 18 where K_y is now in degrees per ft/sec². For this simple case it shows the following:

1. K_ψ must be greater than 0.4487 and less than 7.718 for stability with $K_y = 0$. This can also be inferred from root locus sketches; remember that the end-to-end rate gain is fixed.
2. For K_ψ outside this range, acceleration feedback can stabilize the system; when $K_\psi > 7.718$, large K_y values are required.

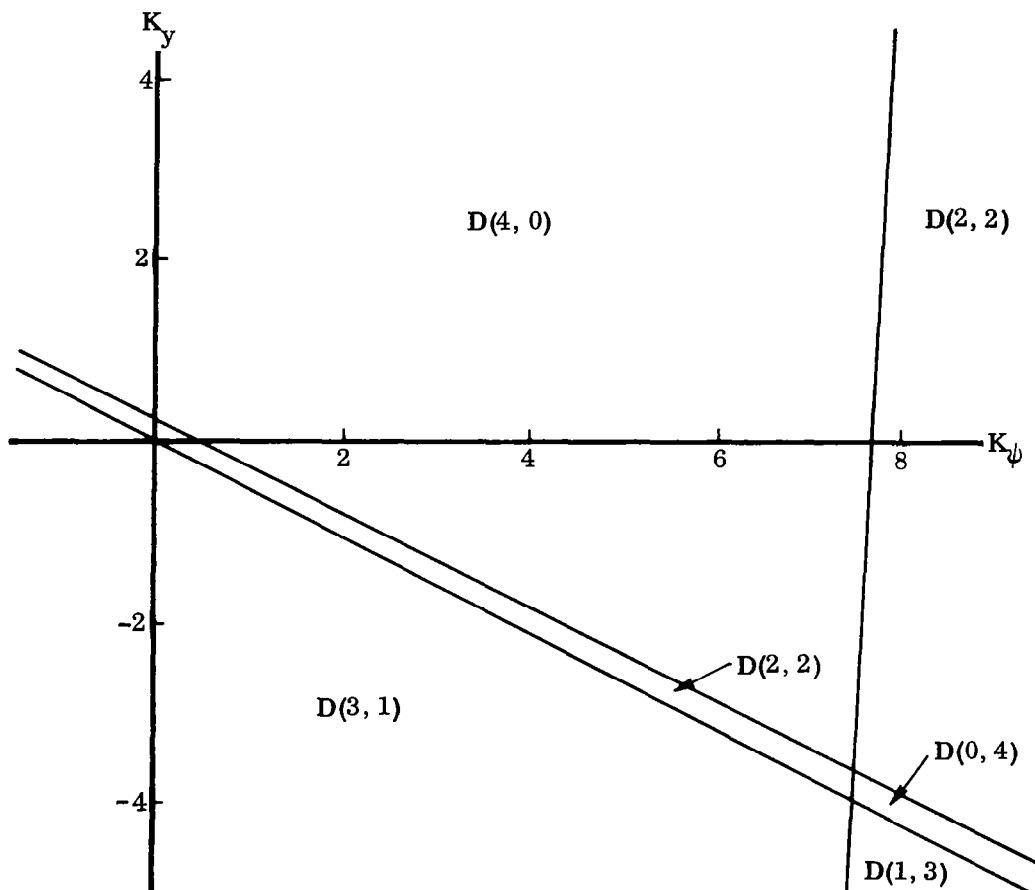


Figure 18. D-Decomposition of K_ψ, K_y Plane for Acceleration Feedback

The limitations are also apparent. The diagram does not indicate, for example, the close proximity of the rigid body roots to the accelerometer loop zeros on the $j\omega$ axis at $\pm j 1.677$ rad/sec as K_y becomes large. The direction of approach is from the left-hand plane, and the roots will never go unstable for this simplified example. The approach would have to be modified in this instance by substituting $s = -\zeta\omega \pm j\omega$ in Equation 120 and solving for the K_ψ and K_y values required.

4.2.5 STABILIZATION OF THE FLEXIBLE VEHICLE BENDING MODES. Once the low-frequency requirements for the launch vehicle flight control system have been roughed out, the control analyst's attention is directed toward ensuring satisfactory stability of the higher-frequency parasitic modes associated with airframe flexibility. These modes of oscillation manifest themselves as a series of lightly damped (rarely greater than a few percent) poles and zeros in the transfer function representing vehicle dynamics. The discussion in this subsection assumes that the frequency separation between the slosh modes and the flexible modes is sufficiently large that they may be considered separately, although as will be shown in Subsection 4.3 the same techniques can be used for cases where this is not true.

4.2.5.1 Equations of Motion. The equations given below are simplifications of the equations derived in Subsection 4.3. They assume the more complex (from the dynamics point of view) situation where the control forces are obtained from gimbaling a relatively massive thrust chamber. Those vehicles employing secondary injection, jetevators, and the like will have a simpler set of equations with those terms dependent on thrust chamber inertia missing.

In writing the equations it is assumed that the modal parameters (generalized mass, $m^{(i)}$; frequency, $\omega^{(i)}$; normalized modal deflection, $\phi_x^{(i)}$; and slope, $\sigma_x^{(i)}$) are calculated in such a fashion that they are orthogonal normal modes or a close approximation thereto. Elastic and inertial crosscoupling terms between the various degrees of freedom (modes of oscillation) do not appear. The only crosscoupling terms that do are dependent upon external forces, for example the thrust and thrust chamber inertial forces. In addition the influence of aerodynamic forces on the launch vehicle is assumed negligible, a good approximation for vehicles having no aerodynamic surfaces. When these terms are included as in Subsection 4.3 the result is to increase the inherent damping of the structural modes and reduce their frequencies by a few percent. These effects are not significant in the "first cut" type of calculations considered here. Certain other crosscoupling terms are neglected, for instance those terms due to thrust chamber inertia (although the effect of inertial loads is included on each mode).

Figure 19 illustrates the yaw plane coordinate system for the launch vehicle. Aerodynamic forces on the rigid vehicle are neglected, leading to significant error in the rigid body control roots. However, the interest here is in the higher-order roots of the system and no attempt is made to retain accuracy over the entire frequency range.

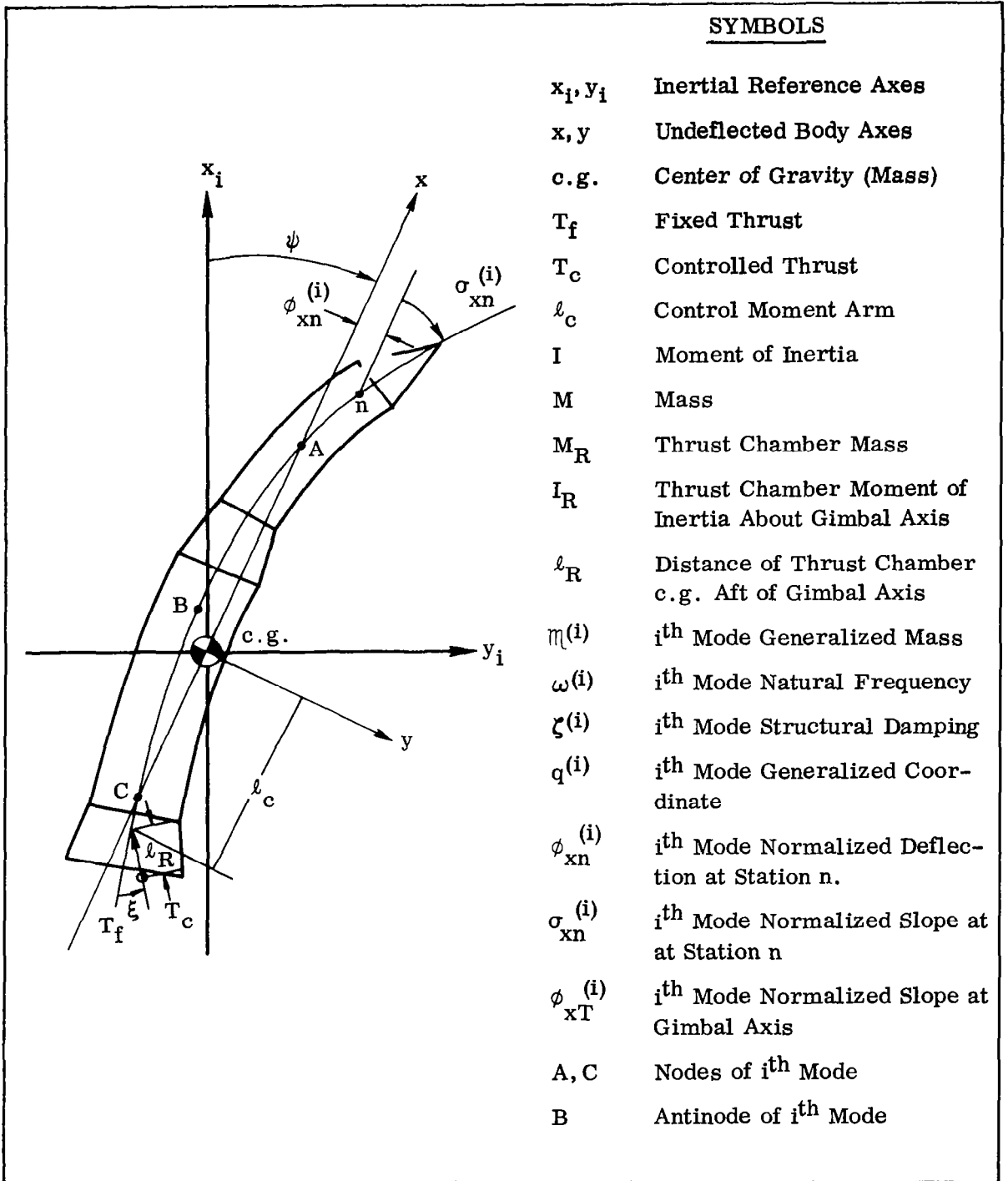


Figure 19. Flexible Vehicle, Yaw Plane Coordinates

The equations of motion are as follows:

Yaw Moment

$$\begin{aligned} I \ddot{\psi} = & T_c \ell_c \left(\xi - \sum_{i=1}^n \sigma_{xT}^{(i)} q^{(i)} \right) - T_f \ell_c \sum_{i=1}^n \sigma_{xT}^{(i)} q^{(i)} \\ & - (T_f + T_c) \sum_{i=1}^n \phi_{xT}^{(i)} q^{(i)} + (M_R \ell_R \ell_c + I_R) \ddot{\xi} \end{aligned} \quad (125)$$

kth Bending Mode

$$\begin{aligned} m^{(k)} \ddot{q}^{(k)} + 2 \zeta^{(k)} \omega^{(k)} m^{(k)} \dot{q}^{(k)} + m^{(k)} [\omega^{(k)}]^2 q^{(k)} \\ = -T_c \phi_{xT}^{(k)} \left(\xi - \sum_{\substack{i=1 \\ i \neq k}}^n \sigma_{xT}^{(i)} q^{(i)} \right) + T_f \phi_{xT}^{(k)} \sum_{\substack{i=1 \\ i \neq k}}^n \sigma_{xT}^{(i)} q^{(i)} \\ - \left(M_R \ell_R \phi_{xT}^{(k)} - I_R \sigma_{xT}^{(k)} \right) \ddot{\xi} \end{aligned} \quad (126)$$

Thrust Chamber Deflection

$$\begin{aligned} \left(s^3 + 2 \zeta_{cn} \omega_{cn} s^2 + \omega_{cn}^2 s + K_c \omega_c^2 \right) \xi = K_c \omega_c^2 \xi_c \\ + (s + K_0) \left\{ \left(1 + \frac{M_R \ell_R \ell_c}{I_R} \right) \ddot{\psi} + \sum_{i=1}^n \left(\sigma_{xT}^{(i)} - \frac{M_R \ell_R \phi_{xT}^{(i)}}{I_R} \right) \dot{q}^{(i)} \right\} \end{aligned} \quad (127)$$

A term by term comparison of these equations with those given in Subsection 4.3 shows that the side acceleration equation has been deleted, and that crosscoupling terms between the first two equations have been neglected together with a number of other terms related to the engine position in each of the orthogonal modes. Thus these equations can be expected to have some errors at and above the engine resonant frequency. The last two terms on the right-hand side of Equation 127 represent inertial load torques acting to deflect the thrust chamber.

In order to demonstrate some of the essential features of this set of equations, further approximations are made to decouple the equations of motion. In effect the cross-coupling between the equations is neglected, which implies that the resultant set of equations is good only for the low bending frequencies. Given this approximation, it

becomes possible to consider only one flexible mode at a time. If a conventional autopilot employing rate and position feedback is assumed, the feedback signal (neglecting sensor dynamics, good only for the lower bending frequencies) is given by

$$\psi_f = (K_R s + 1)\psi + (K_R \sigma_{RG} s + \sigma_{PG})q \quad (128)$$

where the modal index has been omitted since only one mode is being considered. Substituting in this equation for ψ and q ,

$$\begin{aligned} \frac{\psi_f}{\xi}(s) = & (K_R s + 1) \frac{\mu_\xi}{s^2} \left[1 + \frac{M_R \ell_R \ell_c + I_R}{T_c \ell_c} s^2 \right] \\ & - (K_R \sigma_{RG} s + \sigma_{PG}) \frac{T_c \phi_{xT}}{m} \frac{\left[1 + \frac{M_R \ell_R \phi_{xT} - I_R \sigma_{xT}}{T_c \phi_{xT}} s^2 \right]}{s^2 + 2\zeta \omega s + \omega^2} \end{aligned} \quad (129)$$

where σ_{RG} and σ_{PG} represent the normalized modal slopes at the rate and position reference locations, respectively. One now makes the observation that terms involving I_R are negligible with respect to terms in $M_R \ell_R$ and the so-called "tail-wags-dog" zero can be factored out.

$$\begin{aligned} \frac{\psi_f}{\xi}(s) = & \frac{K_R \mu_\xi \left(1 + \frac{M_R \ell_R}{T_c} \right)}{s^2 (s^2 + 2\zeta \omega s + \omega^2)} \left[\left(s + \frac{1}{K_R} \right) (s^2 + 2\zeta \omega s + \omega^2) \right. \\ & \left. - \left(s + \frac{\sigma_{PG}}{K_R \sigma_{RG}} \right) \frac{T_c \phi_{xT} \sigma_{RG}}{m \mu_\xi} s^2 \right] \end{aligned} \quad (130)$$

The frequency of the tail-wags-dog zero,

$$\omega_{TWD} = \sqrt{\frac{T_c}{M_R \ell_R}} \quad (131)$$

represents the frequency at which the lateral force at the gimbal axis due to thrust forces is exactly balanced by the inertial force of thrust chamber motion. For most launch vehicles, this frequency is much higher than the frequency of rigid body motion, and usually considerably higher than the first one or two flexible modes. In those launch vehicles that do not gimbal a relatively massive thrust chamber the frequency is infinite.

The cubic term in the numerator of Equation 130 represents the bending mode and rate gyro zeros. Since these zeros and the gain of the transfer function depend upon the modal slopes at the sensor locations, one concludes that the choice of sensor location has considerable influence on the stability properties of the flight control system. The same is true of the sensor dynamics which (if included in this expression) produce additional pole-zero pairs in the open-loop transfer function.

4.2.5.2 Initial Choice of Sensor Location and Filter Characteristics. It is difficult to specify the means by which flexible modes for any particular launch vehicle can be stabilized without making some assumptions as to the relative importance of various terms in the equations of motion. In this paragraph, the assumptions made up to this point are assumed sufficiently valid to permit initial "first cut" analysis. In the general case, the assumptions must be critically examined before proceeding. The techniques given are therefore only examples; however, they do serve to indicate the many courses of action that may be taken in the solution of the general control problem.

The response of the launch vehicle at high frequencies is usually dominated by the lowest-frequency flexible modes as these have the least inherent damping, the greatest gain through the flight control system, and are the closest in frequency to the atmospheric disturbances that are the potential source of their excitation. Vehicle vibration in one or more of these modes contributes to the bending moment experienced by the vehicle and constitutes a significant contribution to the total vehicle loads, especially for the largest class of launch vehicles where the excitation is greatest. For these reasons the stabilization of these modes is usually achieved by ensuring that the phase of the parasitic signal picked up by the system sensors will not cause regenerative oscillations. It is even possible to phase this signal such that the thrust vector deflection acts to stabilize the oscillation; the closed-loop flexible mode root will have greater damping than the open-loop pole, having a beneficial effect on the vehicle loads.

The higher-frequency flexible modes, by contrast, have greater inherent damping, contribute less to the vehicle loads, and have less gain through the flight control system loop. These modes are gain stabilized as opposed to the phase stabilization of the lower modes; the bending oscillations are decoupled from the flight control system by high attenuation at these higher bending frequencies. In addition the precision of the modal data at the higher modal frequencies makes reliance on phase stabilization speculative; the uncertainties in the mathematical model are too great.

Those flexible modes in the range of the engine position servo resonance are a special case. There may be sufficient friction in the engine gimbal to ensure satisfactory stability. If not, then these modes must be examined closely with the possibility in mind of modifying the position servo transfer function itself to obtain the required degree of stability. In particular the load torques become quite important.

The philosophy used for these examples is therefore based on phase stabilization of the lowest frequency modes and gain stabilization of the higher. Initial choice of

system parameters is based on these two requirements. The importance of the changes in system parameters which take place during a particular stage of flight cannot be overemphasized. If one relies in part on sensor location it must be realized that the modal parameters change significantly during flight. It may be possible to rely on a single sensor location. More probably, some parameters of the system will have to be changed during flight to ensure satisfactory stability. The best design is the one that satisfies the stability criteria used while requiring the fewest inflight changes in the flight control system parameters.

The first consideration is the influence of rate and position sensors on the location of the zeros of Equation 130. The location of the latter is usually dictated by physical considerations. For example, it may be the inertial platform of the launch vehicle flight control system and therefore carried on the last stage of a multistage vehicle. This would imply a forward location. If the rate gyro were located at the same point then the location of the bending mode zero depends only on the modal slope at the sensor location. For the first bending mode, this modal slope increases from some negative value at the aft end of the launch vehicle, through zero at the antinode, and on to increasing positive values as one moves further forward to the nose. The corresponding location of the bending zero moves from some point on the $j\omega$ axis (for the moment, modal damping is neglected) below the bending mode pole, through the pole (at the antinode), and on out the $j\omega$ axis to infinity, returning toward the origin on the real axis. It is assumed in this discussion that the modal deflections are all normalized at some positive value at the gimbal axis, i. e., $\phi_{xT} = 1$. The bending zero location is given by

$$s_{\text{zero}} = \pm \sqrt{\frac{\omega^2}{A - 1}} \quad (132)$$

where

$$A = \frac{T_c \phi_{xT} \sigma_{RG}}{m \mu_\xi}$$

Concurrently with the changing zero location the gain of the transfer function is varying as the quantity $1 - A$. The open-loop gain therefore increases as the sensors are moved to increasing distances from the antinode. The phase of the parasitic signal changes as one moves through the antinode.

Consider the pole-zero plot of Equation 130 shown in Figure 20. The first diagram shows the effect of rate and position sensors being slightly aft of the antinode; the second, slightly forward. Root locus departure angles (the direction in which the root locus emanates from the bending pole) are shown assuming flat (zero phase shift) autopilot response between ψ_f and ξ . The aft location results in inherently stable operation for the bending root; vice versa for the forward location. Additional phase lag due to the engine position servo, sensor dynamics, and autopilot filtering (all desirable since

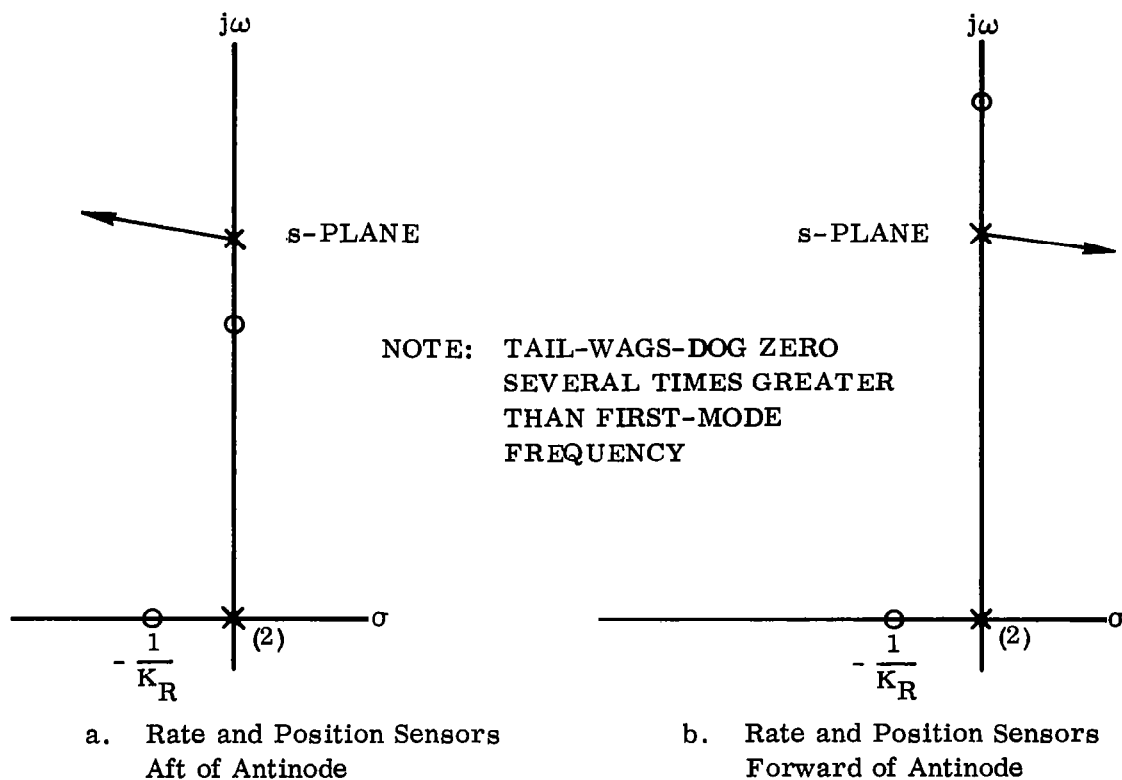


Figure 20. Change in Root Locus Departure Angles with Sensor Location, First Bending Mode

phase lag implies attenuation of the high-frequency bending modes) will cause the direction of departure to rotate clockwise. This benefits the situation shown in the second diagram and deteriorates the first. Further, the modal frequency increases with flight time, meaning increasing amounts of phase lag at the bending frequency. A choice of configurations results. One can choose an aft location with minimal autopilot lag, in which case the critical flight time is toward the end of a particular stage of flight (the locus tends to depart upward, a marginal situation if there is some uncertainty in modal frequency). The other choice is a forward location with sufficient lag for stabilization at the beginning of a particular stage of flight (the locus departs more or less downward, increasing the modal stability as the flight progresses). The aft location implies minimal system lag and attenuation, good for the low (rigid body and propellant slosh) frequencies, bad for the high (higher bending frequencies). The forward choice of sensor location implies the reverse. The higher-frequency modes may enjoy considerable stability margin and closed-loop modal damping while the rigid body and propellant slosh stability suffers. One may be forced into the complication of a simultaneous switch from one sensor location and filter configuration to another part way through a particular stage of flight.

A second and more general situation is to have the rate and position references at different locations on the vehicle. This could arise if the modal slope, hence the gain, at the first-mode frequency is too high (σ_{RG} large) at the position sensor location. In this case both the rate gyro zero and the bending zeros will move as a function of σ_{RG} , assuming σ_{PG} fixed. The same remarks concerning departure angles apply but in this case the bending zeros do not stay on the imaginary axis and may go into the right-hand s-plane. In short, the influence of the position sensor location on the system stability is relatively small. In fact, Equation 130 can be approximated by ignoring the position reference altogether. This amounts to assuming that the phase lead at bending frequencies due to the rate gyro zero is 90 degrees. In this approximation no attempt is made to preserve the proper root locations at rigid body frequencies.

A third possibility is that of using two rate gyros, the so-called "gyro mixing" technique for stabilization of bending modes. In this instance, $K_R \sigma_{RG}$ in Equation 130 is replaced by $K_{R1} \sigma_{RG1} + K_{R2} \sigma_{RG2}$. This offers an additional degree of freedom to the control engineer as the gains can be scheduled to hold a fixed equivalent gain, phase, or some combination of both at bending frequencies. This solution might also be feasible in those situations where it is physically inconvenient to locate a single gyro, for example in a location that suffers from high local levels of vibration. And finally it may be possible to phase stabilize two bending modes at once, which is generally beyond the capabilities of a single rate gyro at all flight times.

The initial choice of autopilot filter is made in concert with the choice of sensor location. The objective is to provide proper phase shift at the frequencies of the bending modes being phase stabilized while maximizing attenuation at high frequencies and minimizing the phase lag at low (below first mode) frequencies. What constitutes proper phase shift at bending frequencies depends on the phase margins desired, which in turn is dependent on not only likely dispersions in autopilot transfer functions but also on analytical uncertainties in the modal parameters. Fortunately, the latter uncertainties are relatively small for the first one or two bending modes unless these modes are intermingled with propellant sloshing (in which case the mode shapes are critically dependent on slosh frequency, in turn a function of thrust and tanking levels and subject to some dispersion). It might arbitrarily be decided that 30 degrees of phase margin are required, and let the initial choice of system configuration be based on this. Later sensitivity studies would determine if this choice of margin is adequate.

It is sometimes feasible to employ lateral accelerometers as a means of stabilizing certain parasitic modes. Considering acceleration feedback alone, the feedback signal is given by

$$\psi_a = K_a (\ell_a \ddot{\psi} + \ddot{y} + \phi_{xa} \ddot{q}) \quad (133)$$

Following the same procedure as before,

$$\frac{\psi_a}{\xi}(s) = K_a(\ell_a - \ell_{cp})(1 - B)\mu_\xi \frac{\left(1 + \frac{M_R \ell_R}{T_c}\right) \left(s^2 + \frac{2\zeta\omega}{1-B}s + \frac{\omega^2}{1-B}\right)}{s^2 + 2\zeta\omega s + \omega^2} \quad (134)$$

where

$$B = \frac{T_c \phi_{xT} \phi_{xa}}{m \mu_\xi}$$

Here again, a pole-zero plot shows a bending mode dipole with the zero below the pole if the normalized modal deflection ϕ_{xa} is negative; the zero is above the pole or on the real axis if ϕ_{xa} is positive. If the accelerometer loop is closed through a pure gain the root locus will depart directly toward the zero. If lag is incorporated due to accelerometer dynamics, autopilot filtering, or servo lags, the root will move toward the right-half plane if the zero is above the pole or on the real axis. (It is assumed that $\ell_a > \ell_{cp}$, a condition for rigid body stability.) One concludes that ϕ_{xa} must in general be negative for stable operation, or if positive there must be more than 180 degrees of phase shift. Given one or another of these conditions the closing of the accelerometer loop will move the root into the left-half s-plane, giving what amounts to artificial damping to the bending mode. Then the rate and position loops can be closed in the usual fashion with considerably more latitude in the gains and filter characteristics used. It must be recognized, however, that the zeros of the propellant slosh modes are also heavily influenced by the choice of accelerometer location and that any choice of location must be examined for its effect on the sloshing stability.

4.2.5.3 Approximate Root Location Via Expansion About the Natural Frequency. A major question that arises from the preceding discussion is associated with the validity of the approximations made by dropping out the various terms in the equations of motion. It would be useful to have a means by which the effect of these terms could be examined without resorting to the brute force technique of solving the resulting equations for the poles and zeros of the transfer function. This can be done, it turns out, by means of expanding the characteristic equation in a Taylor's series about an assumed solution. The latter is usually taken as the natural frequency of the bending mode for which the root is sought. The validity of the approach rests on the root being reasonably close to the natural frequency, usually a good approximation since the gain at the bending frequencies is low by intent. By evaluating the effect of the various terms in the expansion one can obtain a good idea of their relative importance.

The technique is best explained by means of an example. Consider the set of equations given in Paragraph 4.2.5.1 in which the first three bending modes are considered. To this set add Equation 128 for the feedback signal and the following equation for the autopilot transfer function.

$$\xi_c(s) = - \frac{K_A (s + K_I)}{s} F(s) [\psi_f(s) - \psi_c(s)] \quad (135)$$

where $F(s)$ represents the transfer function of the autopilot filter. These equations can be written in matrix form, $\psi_c(s) = 0$:

$$\begin{bmatrix} 1 & \frac{\partial q^{(1)}}{\partial q^{(2)}} & \frac{\partial q^{(1)}}{\partial q^{(3)}} & 0 & \frac{\partial q^{(1)}}{\partial \xi} & 0 \\ \frac{\partial q^{(2)}}{\partial q^{(1)}} & 1 & \frac{\partial q^{(2)}}{\partial q^{(3)}} & 0 & \frac{\partial q^{(2)}}{\partial \xi} & 0 \\ \frac{\partial q^{(3)}}{\partial q^{(1)}} & \frac{\partial q^{(3)}}{\partial q^{(2)}} & 1 & 0 & \frac{\partial q^{(3)}}{\partial \xi} & 0 \\ \frac{\partial \psi}{\partial q^{(1)}} & \frac{\partial \psi}{\partial q^{(2)}} & \frac{\partial \psi}{\partial q^{(3)}} & 1 & \frac{\partial \psi}{\partial \xi} & 0 \\ \frac{\partial \xi}{\partial q^{(1)}} & \frac{\partial \xi}{\partial q^{(2)}} & \frac{\partial \xi}{\partial q^{(3)}} & \frac{\partial \xi}{\partial \psi} & 1 & \frac{\partial \xi}{\partial \xi_c} \\ \frac{\partial \xi_c}{\partial q^{(1)}} & \frac{\partial \xi_c}{\partial q^{(2)}} & \frac{\partial \xi_c}{\partial q^{(3)}} & \frac{\partial \xi_c}{\partial \psi} & 0 & 1 \end{bmatrix} \begin{bmatrix} q^{(1)} \\ q^{(2)} \\ q^{(3)} \\ \psi \\ \xi \\ \xi_c \end{bmatrix} = 0 \quad (136)$$

Here each of the partial derivatives is a function of the complex frequency s . The determinant of this heavily crosscoupled matrix, when set equal to zero, is the characteristic equation of the system. Thus

$$A_{11} - \frac{\partial q^{(1)}}{\partial q^{(2)}} A_{12} + \frac{\partial q^{(1)}}{\partial q^{(3)}} A_{13} + \frac{\partial q^{(1)}}{\partial \xi} A_{15} = 0 \quad (137)$$

where the A's are the minors of the elements in the first row; and we are seeking a solution for the first-mode closed-loop root. This equation may be rewritten,

$$\left[s^2 + 2 \zeta^{(1)} \omega^{(1)} s + \left(\omega^{(1)} \right)^2 \right] = \frac{-1}{\eta^{(1)} A_{11}} \left[- \frac{\partial Q^{(1)}}{\partial q^{(3)}} A_{12} + \frac{\partial Q^{(1)}}{\partial q^{(3)}} A_{13} + \frac{\partial Q^{(1)}}{\partial \xi} A_{15} \right] \quad (138)$$

where $Q^{(1)}$ is the right-hand side of Equation 126 (each partial derivative is a coefficient in s of a term on the right-hand side of this equation). This can be rewritten as

$$\left[s^2 + 2\zeta^{(1)}\omega^{(1)}s + (\omega^{(1)})^2 \right] = \frac{1}{m^{(1)}}G(s) \quad (139)$$

We now wish to expand this equation about the point

$$s_1 = -\zeta^{(1)}\omega^{(1)} + j\omega^{(1)}\sqrt{1 - \zeta^{(1)2}} \cong -\zeta^{(1)}\omega^{(1)} + j\omega^{(1)} \quad (140)$$

One obtains

$$j2\omega^{(1)}\Delta s_1 + \Delta s_1^2 = \frac{1}{m^{(1)}} \left[G(s_0) + G'(s_1)\Delta s_1 + \frac{1}{2!}G''(s_1)(\Delta s_1)^2 + \dots \right] \quad (141)$$

Linearizing,

$$\Delta s_1 = \frac{-jG(s_1)}{2m^{(1)}\omega^{(1)} \left[1 + \frac{jG'(s)}{2m^{(1)}\omega^{(1)}} \right]}$$

Since the bending mode roots are lightly damped,

$$\left. \begin{aligned} G(s) &\cong G(j\omega^{(1)}) + G'(j\omega^{(1)})(-\zeta^{(1)}\omega^{(1)}) \\ G'(s) &\cong G'(j\omega^{(1)}) \end{aligned} \right\} \quad (142)$$

Therefore

$$\Delta s_1 = \frac{-j \left[G(j\omega^{(1)}) - \zeta^{(1)}\omega^{(1)}G'(j\omega^{(1)}) \right]}{2m^{(1)}\omega^{(1)} \left[1 + \frac{jG'(j\omega^{(1)})}{2m^{(1)}\omega^{(1)}} \right]} \quad (143)$$

The terms containing $G'(j\omega^{(1)})$ can usually be neglected except in those cases where there is known to be another root in the near vicinity of the root being sought. In this case it is sufficient to calculate $G'(j\omega^{(1)})$ according to the approximate expression for the total derivative.

$$G'(j\omega^{(1)}) \cong \frac{G[j(\omega^{(1)} + \Delta\omega)] - G(j\omega^{(1)})}{j\Delta\omega} \quad (144)$$

where $\Delta\omega$ is a small increment in the frequency $G(j\omega^{(1)})$, and where $G'(j\omega^{(1)})$ may be evaluated from the matrix by substituting numerical values for the various partial derivatives. By deleting various crosscoupling terms and comparing resultant answers, one can obtain a "feel" for the consequences of such approximations.

This method can, of course, be extended to more complex sets of equations. It turns out to be quite accurate, giving errors in the location of the closed-loop roots within less than 5 percent when compared with machine solutions of the full set of equations. In fact, this method has been used successfully where machine solutions have failed because of ill-conditioned characteristic equations. The accuracy can even be improved if one expands about a point closer to the system root, requiring some prior knowledge or perhaps one iterative cycle on the scheme.

The approximations limit the applicability of the technique to system roots close to the $j\omega$ axis and a relatively short distance from the natural frequency of the bending root. These restrictions can be removed at the cost of some complication in obtaining an answer, necessitating machine solution.

4.2.5.4 Example Case - First-Mode Bending Stability of Liftoff. To illustrate the techniques discussed in the preceding paragraphs, the launch vehicle analyzed earlier for rigid body stability will be examined for first-mode stability at its most critical time, immediately after liftoff. Both the simplified approach wherein most modal crosscoupling and load torque feedback is neglected and the more complex approach outlined in Paragraph 4.2.5.3 will be used for comparison.

The vehicle employs an inertial guidance system on its second stage, thus locating the position reference. The rate reference is unspecified, as are the filter characteristics of the autopilot. The initial choice of gains has been made per the earlier discussion of rigid body stability at maximum dynamic pressure; a conventional (position and rate reference only) autopilot is assumed. Basic data of the configuration are given in Table 5.

To minimize loads caused by excitation of the first mode during booster phase the closed-loop root should be so located as to have increased damping over the open-loop pole during the maximum dynamic pressure flight regime. The design approach chosen is to locate the rate reference forward of the antinode and add sufficient flight control system lag to stabilize the first-mode root at liftoff. The furthest forward excursion of the first-mode antinode during booster stage flight is to station 645. Consequently, the initial choice of rate gyro location is at station 600, sufficiently forward to allow for analytical error in determination of the first bending mode shape. The modal slope at the rate gyro location at liftoff is given by

$$\sigma_{RG}^{(1)} = 0.03932 \text{ ft}^{-1}$$

Table 5. Basic Modal Data for Launch Vehicle at Liftoff - Example Case

I	3,960,000 slug-ft ²
M	9407 slugs
T _c	326,000 lb
T _f	57,760 lb
μ_{ξ}	2.916 sec ⁻²
A _x	40.81 ft/sec ²
ℓ_c	35.40 ft
x _{cg}	787.16 in.
M _R	30.80 slugs
ℓ_R	2.52 ft
I _R	377 slug-ft ²
$\omega^{(1)}$	14.07 rad/sec
$m^{(1)}$	2769.5 slugs
$\phi_{xT}^{(1)}$	1.000 ft/ft
$\sigma_{xT}^{(1)}$	-0.0558 ft ⁻¹
$\sigma_{PG}^{(1)}$	0.0732 ft ⁻¹
antinode at SN 728 inches	

Substituting data from Table 5 into Equation 129 we obtain ($K_R = 0.45$ sec)

$$\frac{\psi_f}{\xi}(s) = \frac{-0.0001846(s^2 + 64.8^2)(s + 2.147)(s + 24.02)(s - 13.46)}{s^2(s^2 + 14.07^2)} \quad (145)$$

The bending mode zeros are located on the positive and negative real axis. With no further lag in the autopilot loop the departure angle of the root locus is at an angle of 24.67 degrees above the horizontal into the right-half plane. Using the simple lag model of the actuator ($\tau_c = 0.054$ sec), one obtains 37.22 degrees of lag at 14.07 radians. The autopilot integrator ($K_I = 0.185 \text{ sec}^{-1}$) gives a 0.75-degree lag. Hence 76.7 degrees of additional lag is required from the autopilot filter to result in the locus departing directly downward.

To obtain the required degree of stability in the presence of autopilot lag variations and to provide damping of the closed-loop root the phase margin is arbitrarily chosen (for this first cut analysis) as 30 degrees. The filter lag requirement at 14.07 rad/sec is therefore 106.7 degrees.

Since the closed-loop root will continue rotating clockwise about the open-loop pole it would be desirable to have the phase change rapidly at the first-mode frequency. A 14.00 rad/sec, 0.5 damped quadratic lag in series with a single lag, $\tau = 0.025$ sec, is picked. The phase change with frequency is more rapid if the damping on the quadratic is lessened, but at the risk of unduly increasing the response at the filter frequencies. This filter choice yields 110 degrees of lag at liftoff. At maximum dynamic pressure the first-mode frequency has increased to almost 17 rad/sec, an increase in lag of about 30 degrees due to filter and engine position servo lags. The tentative choice for the autopilot filter is therefore

$$F(s) = \frac{7840}{(s + 40)(s^2 + 14s + 196)} \quad (146)$$

In this approximate synthesis the following factors have been neglected or approximated:

1. Higher-order engine position servo effects, including inertial loads of the engine.
2. Rate gyro dynamics.
3. Tail-wags-dog zero.
4. Higher-order modes and modal crosscoupling terms.
5. Rigid body plunging and propellant sloshing degrees of freedom.

In addition, this filter configuration should be evaluated for the higher-order bending modes at liftoff to see if sufficient attenuation has been incorporated in the system, as well as for all modes at subsequent flight times to see if this filter is adequate for all flight times. Before refining this "first cut," consideration must be given to the selection of the higher-order dynamic characteristics.

4.2.5.4.1 Selection of Sensor Characteristics. A major consideration in selecting the high-frequency characteristics of the flight control system is that of maintaining close control over variations in the frequency response within the range where such variations have an important bearing on system stability and response properties. The easiest way to do this in the selection of sensor characteristics is to require the frequency response to be more or less flat over this range with minimum gain and phase tolerances. This is particularly true of the rate gyro.

The rate gyros commonly available have underdamped characteristics with a relatively large resonant frequency. The degree of damping is difficult to control within a narrow range (say ± 0.05 about nominal). The natural frequency should therefore be such that the typical variations (± 0.2 about nominal) have little effect on the gyro lag at the highest-frequency mode being phase stabilized. If a range of ± 5 degrees about nominal is a tolerable lag variation at the highest frequency of interest (35 rad/sec) with a damping half of critical, then the gyro frequency should be in excess of 142 rad/sec. Choosing a 0.5-damped, 25-cps (157 rad/sec) gyro seems a reasonable and conservative choice. Variations in the flight control system frequency response within the range of interest to a large extent will be eliminated by this choice.

The position sensor lag is relatively less important at high frequencies. The inertial platform used on this launch vehicle has an essentially flat frequency response within the frequency range where the position reference signal is an important contributor to the system stability properties. Beyond this point its phase and gain are immaterial unless there is pronounced resonant peaking in the response. In this case the contribution of the position sensor dynamics to the overall stability problem would have to be investigated in more detail. For this example it is assumed that such peaking, if it occurs, is of small enough magnitude and high enough frequency that one can assume the response to be flat.

4.2.5.4.2 Selection of Engine Position Servo Characteristics. It would be desirable to use the same philosophy in the selection of the engine position servo response properties as used for the system sensors. One can then rely on closely controlled autopilot filter frequency response. Unfortunately, the large inertia of the thrust chamber limits the degree to which this ideal can be approached. The actuator support structure and the engine inertia determine the servo resonant frequency beyond which the response must roll off according to the approximate relationship (neglecting the contribution of the compressibility of the hydraulic fluid),

$$\omega_c^2 = K_m R^2 / I_R \quad (147)$$

where

K_m = effective spring constant of the actuator and its support structure

R = moment arm of the actuator

I_R = thrust chamber inertia about the gimbal point

For the example under discussion, $\omega_c = 70$ rad/sec. The amplification resulting from the servo resonance must be limited or even eliminated by the proper choice of servo loop gain and damping characteristics such that system oscillations at this frequency will not occur.

The simplest way to accomplish this is to introduce a leakage path for hydraulic fluid between the two sides of the actuator piston and to set the servo loop gain to a relatively low value to minimize excitation of this resonance. The leakage path increases the hydraulic power requirements over that required if no leakage path were used and results in a relatively low static stiffness to applied loads. This leads to some sloppiness in the response at low gimbaling velocities if Coulomb friction in the gimbal bearing is significant. This can cause pronounced nonlinear behavior leading to limit cycles in the engine motion at frequencies depending on the remainder of the launch vehicle dynamics.

A more complex approach is that of using compensation within the servo valve such as derivative (or dynamic) pressure feedback. No leakage orifice is necessary: the static stiffness is higher, resulting in a higher degree of linearity in the response, while the resonant peaking is reduced as before. The increased flexibility in the design (by adjustment of the hydraulic feedback parameters) enables one to use higher servo loop gains as well, thus making increased use of the bandwidth available as limited by the resonant frequency.

Which of these two approaches to use in the general case is a function of system response requirements versus limitations imposed by the servo resonant frequency. The complex approach is necessary in the more difficult situations where response (dictated by the frequency content of expected guidance commands and atmospheric disturbances) and the need for minimizing limit cycle loads are the controlling requirements. It is desirable in any case to minimize variations in response characteristics due to small servo system parameter variations. However, reliability considerations in less severe situations may indicate the simpler approach of leakage path and lower servo loop gain.

For this example the intent is to approximate the engine position servo characteristics which will fall out of detailed studies of the system. The simplest alternative is chosen because the servo resonant frequency is considerably higher than the lag frequency of 18.5 rad/sec considered tolerable for rigid body control. The damping of the complex roots in the servo system response is to be at least 0.3 (neglecting the beneficial effects of gimbal friction). This determines the size of the leakage orifice. The servo loop gain is so chosen that the real root of the actuator transfer function is at half the closed servo loop resonant frequency. Thus the gain of the system at the resonant frequency will not exceed the gain at zero frequency. The results of this choice are

$$K_c = 23.6 \text{ sec}^{-1}$$

$$K_0 = 67.5 \text{ sec}^{-1}$$

$$\omega = 61.4 \text{ sec}^{-1} \text{ (closed-loop resonant frequency)}$$

The servo transfer function neglecting gimbal friction is given by ($T_L(s)$ is the load torque; see Equation 127)

$$\delta(s) = \frac{115,800 \delta_c(s) - \frac{1}{377} (s + 67.5) T_L(s)}{(s + 30.7)(s^2 + 36.8s + 3770)} \quad (148)$$

This choice results in a lag of 32.5 degrees at the first-mode frequency (neglecting load torque feedback) of 14.07 rad/sec at liftoff, more if gimbal friction is included.

4.2.5.4.3 Approximate Root Location. As an example of the approximate method outlined in Paragraph 4.2.5.3, the first-mode closed-loop root is worked out below. The second bending mode influence as well as the load torque feedback terms are included, but gyro dynamics are neglected. The additional data required for the second mode are given in Table 6.

Table 6. Second-Mode Parameters at Liftoff - Example Case

$\omega^{(1)}$	33.95 rad/sec
$\eta^{(1)}$	3136.9 slugs
$\phi_{xT}^{(1)}$	1.0000 ft/ft
$\sigma_{xT}^{(1)}$	-0.1140 ft ⁻¹
$\sigma_{PG}^{(1)}$	-0.1200 ft ⁻¹
$\sigma_{RG}^{(1)}$	0.03676 ft ⁻¹

The matrix equation is given by

$$\begin{bmatrix} 1 & \frac{\partial q^{(1)}}{\partial q^{(2)}} & 0 & \frac{\partial q^{(1)}}{\partial \xi} & 0 \\ \frac{\partial q^{(2)}}{\partial q^{(1)}} & 1 & 0 & \frac{\partial q^{(2)}}{\partial \xi} & 0 \\ \frac{\partial \psi}{\partial q^{(1)}} & \frac{\partial \psi}{\partial q^{(2)}} & 1 & \frac{\partial \psi}{\partial \xi} & 0 \\ \frac{\partial \xi}{\partial q^{(1)}} & \frac{\partial \xi}{\partial q^{(2)}} & \frac{\partial \xi}{\partial \psi} & 1 & \frac{\partial \xi}{\partial \xi_c} \\ \frac{\partial \xi_c}{\partial q^{(1)}} & \frac{\partial \xi_c}{\partial q^{(2)}} & \frac{\partial \xi_c}{\partial \psi} & 0 & 1 \end{bmatrix} \begin{bmatrix} q^{(1)} \\ q^{(2)} \\ \psi \\ \xi \\ \xi_c \end{bmatrix} = 0 \quad (149)$$

where the various partial derivatives are

$$\begin{aligned}
 \frac{\partial q^{(1)}}{\partial q^{(2)}} &= - \frac{(T_c + T_f) \phi_{xT}^{(1)} \sigma_{xT}^{(2)}}{\eta^{(1)} [s^2 + 2\zeta^{(1)} \omega^{(1)} s + (\omega^{(1)})^2]} \\
 \frac{\partial q^{(1)}}{\partial \xi} &= \frac{T_c \phi_{xT}^{(1)}}{\eta^{(1)}} \left[\frac{1 + \frac{M_R \ell_R \phi_{xT}^{(1)} - I_R \sigma_{xT}^{(1)}}{T_c \phi_{xT}^{(1)}} s^2}{[s^2 + 2\zeta^{(1)} \omega^{(1)} s + (\omega^{(1)})^2]} \right] \\
 \frac{\partial q^{(2)}}{\partial q^{(1)}} &= - \frac{(T_c + T_f) \phi_{xT}^{(2)} \sigma_{xT}^{(1)}}{\eta^{(2)} [s^2 + 2\zeta^{(2)} \omega^{(2)} s + (\omega^{(2)})^2]} \\
 \frac{\partial q^{(2)}}{\partial \xi} &= \frac{T_c \phi_{xT}^{(2)}}{\eta^{(2)}} \left[\frac{1 + \frac{M_R \ell_R \phi_{xT}^{(2)} - I_R \sigma_{xT}^{(2)}}{T_c \phi_{xT}^{(2)}} s^2}{[s^2 + 2\zeta^{(2)} \omega^{(2)} s + (\omega^{(2)})^2]} \right] \\
 \frac{\partial \psi}{\partial q^{(1)}} &= \frac{1}{s} \left[\left(\mu_\delta + \frac{T_f \ell_c}{I} \right) \sigma_{xT}^{(1)} + \frac{T_f + T_c}{I} \phi_{xT}^{(1)} \right] \\
 \frac{\partial \psi}{\partial q^{(2)}} &= \frac{1}{s} \left[\left(\mu_\delta + \frac{T_f \ell_c}{I} \right) \sigma_{xT}^{(2)} + \frac{T_f + T_c}{I} \phi_{xT}^{(2)} \right] \\
 \frac{\partial \psi}{\partial \xi} &= - \frac{\mu_\delta}{s} \left(1 + \frac{M_R \ell_R \ell_c + I_R}{T_c \ell_c} s^2 \right) \\
 \frac{\partial \xi}{\partial q^{(1)}} &= - \frac{(s + K_0) \left(\sigma_{xT}^{(1)} - \frac{M_R \ell_R}{I_R} \phi_{xT}^{(1)} \right) s^2}{s^3 + 2\zeta_{cn} \omega_{cn} s^2 + \omega_{cn}^2 s + K_c \omega_c^2}
 \end{aligned} \tag{150}$$

$$\frac{\partial \xi}{\partial q^{(2)}} = - \frac{(s + K_0) \left(\sigma_{xT}^{(2)} - \frac{M_R \ell_R}{I_R} \phi_{xT}^{(2)} \right) s^2}{s^3 + 2 \zeta_{cn} \omega_{cn} s^2 + \omega_{cn}^2 s + K_c \omega_c^2}$$

$$\frac{\partial \xi}{\partial \psi} = - \frac{(s + K_0) \left(1 + \frac{M_R \ell_R \ell_c}{I_R} \right) s^2}{s^3 + 2 \zeta_{cn} \omega_{cn} s^2 + \omega_{cn}^2 s + K_c \omega_c^2}$$

$$\frac{\partial \xi}{\partial \xi_c} = - \frac{K_c \omega_c^2}{s^3 + 2 \zeta_{cn} \omega_{cn} s^2 + \omega_{cn}^2 s + K_c \omega_c^2}$$

$$\frac{\partial \xi_c}{\partial q^{(1)}} = \frac{K_A K_R (s + K_I) \left(\sigma_{RG}^{(1)} s + \frac{\sigma_{PG}^{(1)}}{K_R} \right)}{s (\tau_f s + 1) \left(\frac{s^2}{\omega_f^2} + \frac{2 \zeta_f}{\omega_f} s + 1 \right)}$$

$$\frac{\partial \xi_c}{\partial q^{(2)}} = \frac{K_A K_R (s + K_I) \left(\sigma_{RG}^{(2)} s + \frac{\sigma_{PG}^{(2)}}{K_R} \right)}{s (\tau_f s + 1) \left(\frac{s^2}{\omega_f^2} + \frac{2 \zeta_f}{\omega_f} s + 1 \right)}$$

$$\frac{\partial \xi_c}{\partial \psi} = \frac{K_A K_R (s + K_I) \left(s + \frac{1}{K_R} \right)}{s (\tau_f s + 1) \left(\frac{s^2}{\omega_f^2} + \frac{2 \zeta_f}{\omega_f} s + 1 \right)}$$

Evaluating at $s = j 14.07$ the matrix becomes (using the usual notation for magnitude and phase)

$$\begin{bmatrix} s^2 + 2\zeta^{(1)} \omega^{(1)} s + (\omega^{(1)})^2 & 15.8 & 0 & 110.6 & 0 \\ 0.00716 & 1 & 0 & 0.1011 & 0 \\ 0.000476 & 0.001491 & 1 & 0.01395 & 0 \\ 0.0294/\underline{157.55^\circ} & 0.0361/\underline{157.55^\circ} & 0.937/\underline{157.55^\circ} & 1 & 0.903/\underline{147.1^\circ} \\ 0.426/\underline{-37.11^\circ} & 0.431/\underline{6.63^\circ} & 10.40/\underline{-29.67^\circ} & 0 & 1 \end{bmatrix} \quad (151)$$

from which are obtained

$$\left. \begin{aligned} A_{11} &= 0.932/\underline{7.78^\circ} \\ A_{12} &= 0.3589/\underline{95.7^\circ} \\ A_{14} &= 0.360/\underline{-73.81^\circ} \end{aligned} \right\} \quad (152)$$

and

$$\begin{aligned} \frac{G(j 14.07)}{m_l^{(1)}} &= \frac{1}{0.932/\underline{7.78^\circ}} \left[15.8 \times 0.03589/\underline{95.7^\circ} + 110.6 \times 0.360/\underline{-73.81^\circ} \right] \\ &= 42.2/\underline{-81.44^\circ} \end{aligned} \quad (153)$$

from which it is seen that the second-mode contribution is small. Using

$$\Delta s_1 = \frac{-j G(j\omega^{(1)})}{2m_l^{(1)} \omega^{(1)}} \quad (154)$$

the following result is obtained.

$$\Delta s_1 = -1.48 - j 0.2232 \quad (155)$$

Thus the first-mode closed-loop root is at

$$s^{(1)} = -1.48 \pm j 13.85 \quad (156)$$

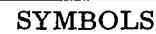
The increased lag of the root is attributed primarily to the effect of load torque feedback. The fact that the root does not move radially from the pole as the gain is increased, but curves upward, means that damping of slightly more than ten percent of critical has been achieved.

4.2.6 STABILIZATION OF PROPELLANT SLOSH MODES. After the high-frequency characteristics of the flight control system have been estimated, sloshing stability is analyzed. The rigid body modes are included because of their proximity to the slosh modes.

4.2.6.1 Equations of Motion. Because of the symmetry of most launch vehicles, analysis of one control axis (motion in one plane) usually suffices. The vehicle is considered a rigid airframe with propellant tanks containing fluid that is free to slosh. Solution of the hydrodynamic equations describing slosh motion provides answers in the form of forces and moments acting on the airframe, which can be represented by an equivalent simple pendulum. The mathematical model thus consists of the airframe with a series of simple pendulums representing sloshing dynamics.

Figure 21 illustrates the coordinate system used in analyzing propellant sloshing stability. The m propellant pendulums are hinged at various points on the vehicle centerline and the assumption is made that $\beta = \psi$, usually true within the frequency range of propellant slosh. For the present discussion the propellant slosh damping is assumed to be zero. The equations of motion can be rewritten as

$$\begin{bmatrix}
 (s^2 - \mu_\beta) & 0 & \mu_1 & \mu_2 & \dots & \mu_m \\
 -\frac{F_\beta}{M_0} & 1 & \frac{A_x}{M_0} M_1 & \frac{A_x}{M_0} M_2 & \dots & \frac{A_x}{M_0} M_m \\
 -\frac{\ell_1 - L_1}{L_1} s^2 & -\frac{1}{L_1} & s^2 + \omega_1^2 & 0 & \dots & 0 \\
 -\frac{\ell_2 - L_2}{L_2} s^2 & -\frac{1}{L_2} & 0 & s^2 + \omega_2^2 & \dots & \vdots \\
 \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
 -\frac{\ell_m - L_m}{L_m} s^2 & -\frac{1}{L_m} & 0 & 0 & \dots & s^2 + \omega_m^2
 \end{bmatrix}
 \begin{bmatrix}
 \psi \\
 \ddot{y} \\
 \beta_1 \\
 \vdots \\
 \beta_m
 \end{bmatrix}
 =
 \begin{bmatrix}
 \mu_\xi \\
 -\frac{T_c}{M_0} \\
 0 \\
 \vdots \\
 0
 \end{bmatrix}
 \xi
 \quad (157)$$



x_i, y_i	Inertial Reference Axes
x, y	Body Axes
c.p.	Center of Aerodynamic Pressure
c.g.	Reduced Center-of-Gravity (Mass)
T_f	Fixed Thrust
T_c	Controlled Thrust
ℓ_β	Aerodynamic Moment Arm
ℓ_c	Control Moment Arm
ℓ_i	i^{th} Pendulum Moment Arm
L_i	i^{th} Pendulum Length
M_i	i^{th} Pendulum Mass
F_β	Lateral Aerodynamic Force Coefficient
ξ	Controlled Thrust Vector Angle
β_i	i^{th} Pendulum Angle
ψ	Yaw Angle
I_0	Reduced Moment of Inertia
M_0	Reduced Mass
M	Total Mass (Equals $M_0 + \sum_{i=1}^n M_i$)
μ_ξ	= $T_c \ell_c / I_0$, Control Effectiveness Parameter
μ_β	= $F_\beta \ell_R / I_0$, Aerodynamic Instability Parameter
μ_i	= $M_i \ell_i A_x / I_0$
ω_i	= A_x / L_i , Slosh Frequency, i^{th} Pendulum

EQUATIONS OF MOTION

Yaw Moment:

$$I_0 \ddot{\psi} = T_c \ell_c \xi + F_\beta \ell_\beta \psi - \sum_{i=1}^n M_i \ell_i A_x \beta_i$$

Side Acceleration:

$$M_0 \ddot{y} = -T_c \xi + F_\beta \psi - \sum_{i=1}^n M_i A_x \beta_i$$

Pendulum Angle:

$$L_i \ddot{\beta}_i + A_x \ddot{\beta}_i = \ddot{y} + (\ell_i - L_i) \ddot{\psi}$$

Axial Acceleration:

$$MA_x = T_c + T_f - D$$

Figure 21. Rigid Vehicle with Propellant Sloshing, Yaw Plane Coordinates and Equations of Motion

The transfer function relating yaw angle ψ to control thrust vector deflection angle ξ is given by

$$\begin{aligned}
& \left\{ s^2 \left[1 + \sum_{i=1}^m \frac{M_i}{M_0} \left(1 + \frac{\ell_i(\ell_i - L_i)}{r^2} \right) \frac{\omega_i^2}{s^2 + \omega_i^2} \right. \right. \\
& \quad \left. \left. + \sum_{i=1}^m \frac{M_i}{M_0} \frac{\ell_i}{r^2} \frac{\omega_i^2}{s^2 + \omega_i^2} \sum_{\substack{j=1 \\ j \neq i}}^m \frac{M_j}{M_0} (\ell_i - L_i - \ell_j + L_j) \frac{\omega_j^2}{s^2 + \omega_j^2} \right] \right. \\
& \quad \left. - \mu_\beta \left[1 + \sum_{i=1}^m \frac{M_i}{M_0} \left(1 - \frac{\ell_i}{\ell_\beta} \right) \frac{\omega_i^2}{s^2 + \omega_i^2} \right] \right\} \psi(s) \\
& = \mu_\xi \left[1 + \sum_{i=1}^m \frac{M_i}{M_0} \left(1 + \frac{\ell_i}{\ell_c} \right) \frac{\omega_i^2}{s^2 + \omega_i^2} \right] \xi(s) \tag{158}
\end{aligned}$$

where $r = \sqrt{I_0/M_0}$, the radius of gyration.

This rather formidable expression indicates that the presence of the pendulums alter the rigid body characteristics and vice versa. In particular:

1. The second term represents the effect of one pendulum on another and is zero for any set of two pendulums only if the pendulum masses or hinge points are located at the same point on the vehicle. Normally the terms are small because $M_i/M_0 \ll 1$, whence the products of these factors become very small. Note that slosh pendulums influencing each other most strongly are located on opposite ends of the vehicle, resulting in the greatest magnitude for these terms.
2. The form of the third term shows that the vehicle aerodynamic instability parameter, μ_β , is altered by the presence of the slosh pendulums. Ignoring the second term for the moment, it is seen that the term $s^2 - \mu_\beta$ can be factored out only if the pendulum masses are located at the center of percussion with respect to the center of aerodynamic pressure, i.e., $r^2 + \ell_\beta(\ell_i - L_i) = 0$.
3. Assuming negligible aerodynamics ($\mu_\beta = 0$) and crosscoupling between pendulums ($M_i/M_0 \ll 1$), the numerator of $\psi/\xi(s)$ is identical to the denominator if the pendulum masses are located at the center of percussion, i.e., $r^2 - \ell_c(\ell_i - L_i) = 0$.

To factor Equation 158 without resorting to a digital computer routine, one can employ a root-locus-derived technique. Note that the polynomials representing numerator and denominator are polynomials in s^2 . Thus all poles and zeros lie on either the real axis, the imaginary axis, or possibly symmetrically placed about the origin. Substituting $z = s^2$ and taking the zeros of the transfer function (the right-hand side of Equation 131) one has an equation of the form,

$$1 + \sum_{i=1}^m \frac{A_i}{z + a_i} = \frac{\prod_{i=1}^m (\text{zeros})}{\prod_{i=1}^m (\text{poles})} \quad (159)$$

where

$$A_i = \frac{M_i}{M_0} \left(1 + \frac{\ell_i}{\ell_c} \right) \omega_i^2$$

$$a_i = \omega_i^2$$

This can be solved by adding one mode at a time in a manner best explained by referring to Figure 22. Since the locus generally lies on the real axis, only the gain criteria need be satisfied to determine successively b_1 , c_1 , and c_2 and finally d_1 , d_2 , and d_3 in this example. For each step the poles become the roots obtained from the preceding step with the addition of the pole of the new mode being added; the zeros are the poles of the preceding step. The gain, of course, is the A_i pertinent to each step in the process.

For the slosh poles the situation is more complex but in principle it is the same. The factors of the coefficients of $-\mu_\beta$ and s^2 (neglecting the crosscoupling terms) on the left-hand side of Equation 158 are obtained using the same technique. To add in the crosscoupling terms, the same procedure is repeated, only in this case two new poles (which cancel two existing zeros) are added. For each successive step the zeros are the a_i and the poles are the roots of the previous step plus the two additional ones canceling two zeros (see Figure 23). Each term in Equation 158 is now a factored polynomial after multiplying through by

$$\prod_{i=1}^n (z + a_i)$$

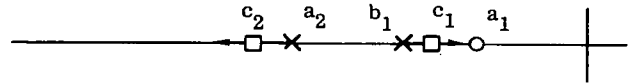
Therefore,

$$\left[z \prod_{i=1}^m (z + f_i) - \mu_\beta \prod_{i=1}^m (z + g_i) \right] \psi(z) = \mu_\xi \prod_{i=1}^m (z + d_i) \xi(z) \quad (160)$$

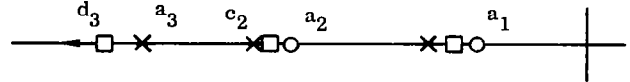
$$1 + \frac{A_1}{z + a_1} = \frac{z + b_1}{z + a_1}$$



$$\frac{z + b_1}{z + a_1} + \frac{A_2}{z + a_2} = \frac{(z + c_1)(z + c_2)}{(z + a_1)(z + a_2)}$$



$$\frac{(z + c_1)(z + c_2)}{(z + a_1)(z + a_2)} + \frac{A_3}{z + a_3} = \frac{(z + d_1)(z + d_2)(z + d_3)}{(z + a_1)(z + a_2)(z + a_3)}$$



FINAL z-PLANE CONFIGURATION

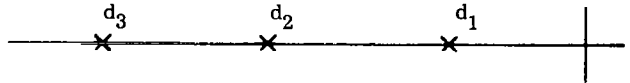


Figure 22. Sloss Zero Location Using Root Locus Technique

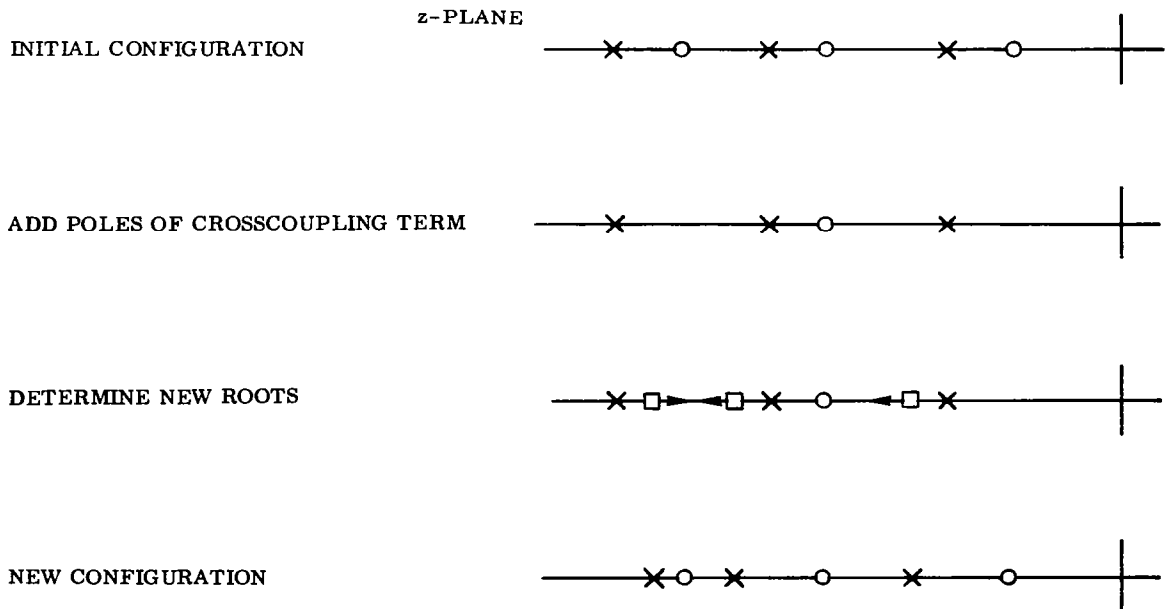


Figure 23. Sloss Pole Shifting Due to Crosscoupling

If the g_i are treated as zeros, the f_i as poles (with an additional one at the origin), and μ_β as a negative gain (the angle criterion is now 0, not 180 degrees), the poles of the transfer function are finally obtained. The square root of each z-plane pole and zero yields the s-plane poles and zeros of $\psi/\xi(s)$.

This procedure, while rather involved, does demonstrate the relative importance of each slosh mode to the overall problem. For example, if one particular mode has a slosh mass M_i considerably smaller than any other, a term approximately equal to $s^2 + \omega_i^2$ will factor out of both numerator and denominator. The result is a dipole of small residue (small pole-to-zero separation) at a frequency closely approximated by the natural slosh frequency ω_i of the particular tank. On the other hand, large slosh masses generally lead to upward shifts in frequency; the poles and zeros appear at frequencies somewhat higher than ω_i and the pole-zero separation becomes more sizeable. Various simplifying assumptions can be made such as neglecting second-order terms in M_i/M_0 and neglecting aerodynamics (see Reference 7). The end result is a series of poles and zeros strung along the $j\omega$ axis, together with two real poles equally spaced from the origin on the real axis.

We are dealing here with the transfer function $\psi/\xi(s)$, which will not have complex zeros (four) symmetrically placed about the origin. This results from the A_i of Equation 159 being positive (no pendulum is hinged aft of the thrust vector gimbal point). A similar argument would hold for the poles except that a massive pendulum forward of the aerodynamic center of pressure could result in complex zeros for the third term in Equation 158, leading (perhaps) to complex poles of the overall transfer function $\psi/\xi(s)$ where μ_β is large. This is not a typical situation as the forwardmost tanks on a multi-stage vehicle generally have the smallest slosh masses. The foregoing argument for the transfer function zeros does not hold if acceleration feedback is used. The zeros of the transfer function relating lateral acceleration at some launch vehicle station to engine deflection can be complex.

Each pole corresponds to an eigenvalue of the matrix of Equation 157; that is, a natural frequency of vibration of the combined system of vehicle and slosh pendulums. With this eigenvalue there is an associated eigenvector, or mode shape, that defines the relative amplitudes of coordinates ψ , y , and β_i when the uncontrolled system is oscillating at this frequency. Each pole therefore represents a mode of oscillation in which all pendulums are moving. If the poles are relatively widely separated and close to the ω_i , each mode is dominated by motion of one pendulum. Under these circumstances the modes can be referred to by the dominant tank, for example the "booster stage fuel tank slosh mode." However, close proximity of frequencies (due to a common tank diameter for several tanks) can result in sizable motion in several tanks and the referral then becomes incorrect. The distinction is important because it may become necessary to stabilize a slosh mode using baffles and the question is then one of which tank or tanks should be baffled.

4.2.6.2 Design Considerations. The pole-zero sequence along the imaginary axis determines the difficulty to be encountered in stabilizing the slosh modes. If the sequence with increasing frequency is zero, pole, zero, pole, etc., the system is relatively easy to stabilize. In the absence of any compensation, closing a feedback path around the transfer function $\frac{\psi}{\xi}(s)$ results in loci departing directly up the imaginary axis from the poles. With the net lead from the combination of autopilot filtering, rate feedback, and actuator lag (the pole-zero dipole at the origin contributes very little lag), the loci depart upward into the left-hand plane. This situation is illustrated by Figure 24a and b.

If one of the dipoles in the pole-zero sequence is reversed the locus emanating from that pole will depart downward and into the right-half plane, creating a closed-loop sloshing instability. This situation is illustrated in Figure 24c. Such a dipole has what is termed an unstable configuration. Since such a dipole requires lag for stabilization (to rotate the locus into the left-hand plane), while the remaining dipoles are destabilized with increasing lag, the situation usually calls for antislosh baffles.

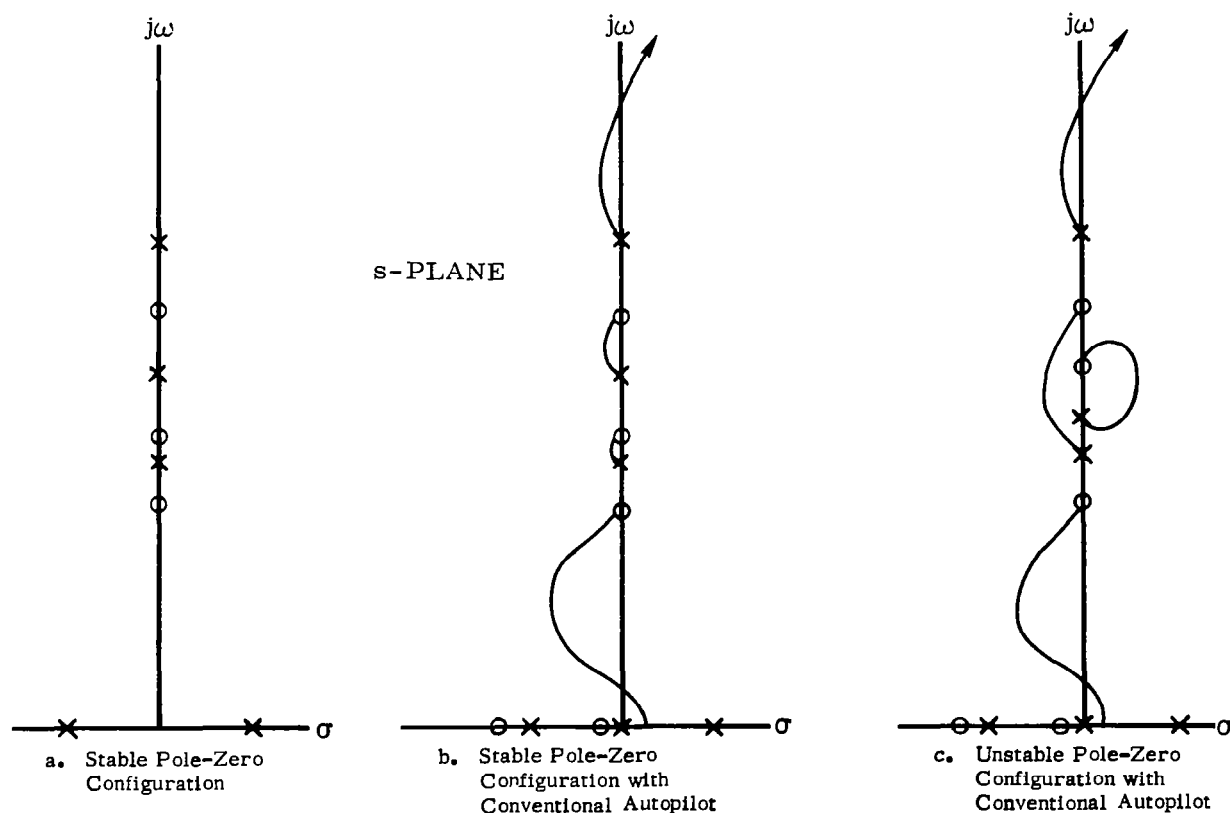


Figure 24. Stable and Unstable Slosh Pole-Zero Configurations

In general the pole-zero sequence has the majority of dipoles in a stable configuration, with the result that autopilot lag tends to destabilize the system. Since the slosh frequencies generally increase with increasing flight time, due to increasing axial acceleration, the root locus departure angles from the poles representing the various slosh modes tend to rotate clockwise into the right-half plane; flight control system lag increases at the higher frequencies. Aggravating this situation is the decreasing vehicle mass which increases vehicle sensitivity to sloshing. The result is a tendency for the sloshing stability problem to be at its worst toward the end of a particular stage of flight.

Up to this point the slosh damping has been ignored. In practice it is usually very small, particularly if the tank has a large free surface area and is "clean" internally, having no internal structural members. The approximation made is quite good unless baffles or other internal structure are added. In this case the poles associated with vehicle slosh modes having considerable motion in the baffled tank tend to move to the left. The damping is nonlinear, increasing with increasing fluid motion. Thus some slosh activity is necessary to provide the damping. If one relies on baffling to stabilize an otherwise marginally stable sloshing mode, a slosh limit cycle is the result. Low fluid motion (low damping) leads to an unstable root which causes a divergent response at the frequency of the unstable root. As the amplitude increases, so does the damping and the open-loop pole and closed-loop root more toward the left. The stable operating point is at some limit cycle amplitude where the damping is sufficient to place the closed-loop root on the imaginary axis. The reverse argument shows that an initially damped closed-loop root will move toward the $j\omega$ axis as the modal response amplitude dies down, reducing the damping. Clearly the amount of baffling determines the limit cycle amplitude.

Slosh limit cycles also arise as a result of the nonlinear nature of electrohydraulic or electropneumatic engine positioning servos (see Paragraph 4.2.7.1). These arise out of the servo lag's dependence on the amplitude of engine motion and will result in stable or unstable limit cycle operating points.

This discussion serves to indicate the consequences of using one or another means for stabilizing propellant sloshing:

1. **Passive Stabilization** - Requires modifying the internal structure of the tank to increase the damping (baffles) or raise the slosh frequency (partitioning).
2. **Active Stabilization** - Requires modifying the flight control system via addition of auxiliary loops (direct slosh pressure feedback or, more commonly, accelerometer feedback) or network shaping.

Either of these approaches carries its own advantages and disadvantages. Active compensation implies additional flight control system elements that must work to ensure mission success. Further, it is not always possible to synthesize such a system.

On the other hand, there is little if any weight penalty. Passive means are more fool proof, but they carry a weight penalty. This is particularly true if baffles or tank partitioning is necessary in the upper stages of a launch vehicle. Tank partitioning also complicates the dynamic model; although modal frequency is raised and slosh mass is reduced, multiple-mode frequencies are produced because the cross-section is no longer cylindrically symmetric. In either case, additional slosh modal damping mitigates inflight slosh loads that will be experienced, important since the frequencies are within the range of atmospheric disturbances.

4.2.6.3 Example Case - Slosh Stability Prior to Staging. The basic data pertaining to the rigid body and propellant sloshing stability at staging are given in Table 7. Several points are immediately apparent:

1. The aerodynamic forces are very low, as the maximum dynamic pressure region has been passed.
2. The fourth pendulum corresponds to a liquid-hydrogen tank on the upper stage of the vehicle; its influence can be ignored as a consequence of $M_4/M_0 = 0.00301$.
3. The control loop gain of $K_A = 1.75$, while adequate for the maximum dynamic pressure region, is probably too high here because of the increase in μ_ξ to a much higher value.

Table 7. "Reduced" Rigid Body and Slosh Parameters for Launch Vehicle Prior to Staging, Example Case

$M_0 = 1998$ slugs	$\omega_1 = 6.6297$ rad/sec
$I_0 = 1,857,000$ slug-ft ²	$M_1 = 190.92$ slugs
$\ell_c = 47.61$ ft	$L_1 = 3.986$ ft
$\ell_\beta = 15.01$ ft	$\ell_1 = -39.985$ ft
$\mu_\xi = 9.539$ sec ⁻²	$\omega_2 = 7.4325$ rad/sec
$\mu_\beta = 0.1064$ sec ⁻²	$M_2 = 331.5$ slugs
$T_c = 375,300$ lb	$L_2 = 3.155$ ft
$A_x = 174.3$ ft/sec ²	$\ell_2 = 20.937$ ft
$\omega_3 = 11.814$ rad/sec	$\omega_4 = 9.735$ rad/sec
$M_3 = 98.16$ slugs	$M_4 = 6.013$ slugs
$L_3 = 1.249$ ft	$L_4 = 1.8395$ ft
$\ell_3 = 19.416$ ft	$\ell_4 = 36.324$ ft
$r^2 = 938.5$ ft ²	

Equation 158 can be rewritten for this application as

$$\begin{aligned}
& s^2 \left[1 + \sum_{i=1}^3 \frac{M_i}{M_0} \left(1 + \frac{\ell_i (\ell_i - L_i)}{r^2} \right) \frac{\omega_i^2}{s^2 + \omega_i^2} \right. \\
& \left. + \sum_{i=1}^3 \frac{M_i}{M_0} \frac{\ell_i}{r} \frac{\omega_i^2}{s^2 + \omega_i^2} \sum_{\substack{j=1 \\ j \neq i}}^3 \frac{M_j}{M_0} (\ell_i - L_i - \ell_j + L_j) \frac{\omega_j^2}{s^2 + \omega_j^2} \right] \psi(s) \\
& = \mu_{\xi} \left[1 + \sum_{i=1}^3 \frac{M_i}{M_0} \left(1 + \frac{\ell_i}{\ell_c} \right) \frac{\omega_i^2}{s^2 + \omega_i^2} \right] \xi(s)
\end{aligned} \tag{161}$$

Evaluating using Table 7 yields ($s^2 = z$)

$$\begin{aligned}
& z \left[1 + \frac{12.065}{z + 43.95} + \frac{14.1}{z + 55.25} + \frac{9.46}{z + 139.8} + \frac{15.54}{(z + 43.95)(z + 55.25)} \right. \\
& \left. + \frac{113.69}{(z + 43.95)(z + 139.8)} + \frac{114.7}{(z + 55.25)(z + 139.8)} \right] \psi(z) \\
& = 9.539 \left[1 + \frac{0.6745}{z + 43.95} + \frac{5.15}{z + 55.25} + \frac{9.67}{z + 139.8} \right] \xi(z)
\end{aligned} \tag{162}$$

When solved for roots the transfer function becomes

$$\frac{\psi}{\xi}(s) = \frac{9.539 (s^2 + 44.08) (s^2 + 60.51) (s^2 + 149.9)}{s^2 (s^2 + 48.50) (s^2 + 76.89) (s^2 + 149.2)} \tag{163}$$

The highest-frequency pole-zero dipole has an unstable configuration; on the other hand, its influence on the stability of the system will be small because of the small residue at the pole (the pole-zero separation distance is less than 0.15 rad/sec). With a small amount of damping (say, one percent) in the tank it will be stable for this flight time.

The lowest-frequency pole-zero dipole has a frequency separation of about 0.3 rad/sec and will have the next smallest influence; the intermediate dipole is separated by about 1.0 rad/sec and will dominate the sloshing stability qualities.

Considering the slosh pole at 8.77 rad/sec, the lag through the flight control system (per Paragraph 4.2.5.4) is 5.6 degrees at this frequency, meaning that the closed-loop root will, in the absence of any damping in the tank, be in the right-half plane. Other than adding baffles to the vehicle, the stability qualities can be improved only by decreasing the filter lag (at this flight time, the first bending mode does not need as

much lag as at liftoff), increasing the rate gain (good for perhaps two or three degrees of lead), and reducing the autopilot gain by half (the μ_{ξ} has doubled from its value at liftoff).

The solution chosen depends upon the means by which the changes in autopilot characteristics are implemented. If, for example, the gains are continuously varied throughout flight with perhaps one or two changes in filter characteristics the gains would be chosen to be as low as possible consistent with requirements for rigid body response. If a minimal number of discrete changes in characteristics is the goal then the gains would be changed as soon as possible, after passage through the maximum dynamic pressure region, to as low a value as possible and the filter to as little lag as possible consistent with the requirements for first bending mode response. The particular vehicle under discussion (Atlas/Centaur, as flown) changes the position gain K_A to less than half its former value three quarters of the way through booster phase. To compensate for marginal rigid body stability at this time (the dynamic pressure is 60% of its maximum value) the rate gain is increased by 25%. In addition baffles are used to stabilize the particular slosh mode under discussion here; and the filter is changed to one of about 10 degrees less phase shift at slosh frequencies at about 30% through the booster phase when the first bending mode frequency has risen sufficiently to permit it.

4.2.7 LIMIT CYCLES. The discussion in the preceding paragraph has pointed out two major sources of limit cycle behavior in the dynamics of launch vehicles, the thrust chamber positioning servo and propellant slosh damping. These arise from unintentional system nonlinearities and necessitate some discussion of how these nonlinearities can be handled in what otherwise is a linear analysis. The intent is only to outline possible procedures as nonlinear analytical techniques are, strictly speaking, outside the scope of this monograph. These procedures fall into two general areas, one analogous to frequency response techniques using one or another form of describing function, the other being direct simulation.

4.2.7.1 Use of Describing Functions. The usual technique employed is that of quasi-linearization by means of describing functions (Reference 2). This approach is restricted to the prediction of possible limit cycles (there is no guarantee that a limit cycle predicted by this technique will in fact exist) in a system having low bandpass frequency response characteristics between the output of the nonlinearity and the input. This permits one to assume that only the fundamental component of the output of the nonlinearity remains after passing through the remainder of the system back to the input. If the input to the nonlinear element is a pure sinusoid the output will be a periodic function of time (true for the class of nonlinearities being considered) which can be stated in the form of a Fourier series. The ratio of the fundamental component of the Fourier series (that term having the same frequency as the input) to the input sinusoid is the describing function for the nonlinear element, a function of both amplitude and frequency of the input sinusoid. Using the describing function, one then examines the

system stability to see if there are signal amplitudes and frequencies at which a "root" of the system is marginally stable. If so, this constitutes an operating point for a possible limit cycle. The stability or instability of the limit cycle is inferred from the motion of the "root" with changes in signal amplitude and frequency.

4.2.7.1.1 Thrust Chamber Positioning Servos. If the launch vehicle employs gimbaled thrust chambers for thrust vector control, a possible source of nonlinear behavior leading to limit cycles is the rocket engine positioning servo. Because of the dependency of engine motion on both engine commanded position and load torque feedback (inertial loads on the thrust chamber due to vehicle motion), the describing function is most easily given in terms of the fundamental component of the output engine motion having a zero-to-peak amplitude of $\bar{\xi}$ (see Reference 8). The major sources of nonlinear behavior within the actuation scheme are Coulomb friction with the gimbal bearing and the nonlinear (square root) flow-pressure relationship for hydraulic fluid flow through an orifice. The transfer function relating engine position to engine command and load torque has coefficients that are functions of the amplitude and frequency of engine motion.

Because of the dependence of the describing function for the positioning servo on frequency as well as amplitude, the analysis of the system for limit cycles is restricted to considering one frequency at a time. In using it for investigating the possible existence of a flexible mode limit cycle, the servo transfer function coefficients are computed for several assumed amplitudes at the frequency corresponding to the bending mode closed-loop root. (In practice the open-loop pole natural frequency is used since the closed-loop root seldom moves so far from the pole to render the describing function seriously in error.) Using this transfer function, one plots for each assumed amplitude the root locus corresponding to the root whose stability is being investigated, or the Nyquist diagram for the range of frequencies close to the closed-loop root, and infers from these plots the stability or instability of the limit cycle. At any other range of frequencies, in the Nyquist plot or any other root locus plot, the results are meaningless because the position servo describing function is invalid. In particular, certain of these other roots may be "unstable." Whether they are in fact unstable requires investigation using the describing function computed for their particular frequencies. For the example case here the root locus plot might look like that shown in Figure 25, the Nyquist diagram in Figure 26. These show a potentially unstable situation. If a disturbance input excited this particular parasitic mode to an engine amplitude exceeding $\bar{\xi} = 0.4$ degree, the mode would continue in divergent oscillations because the increasing engine motion implied by the positive real part of the bending root causes increased phase lead, which further increases the rate of divergence. On the other hand, the same reasoning shows that engine amplitudes less than 0.4 degree will result in decaying oscillations. The 0.4-degree point represents an unstable operating point for the limit cycle. One further point: The open-loop poles (Figure 25) move as well as the closed-loop roots with changing engine amplitude. This is a result of the load torque feedback on engine motion.

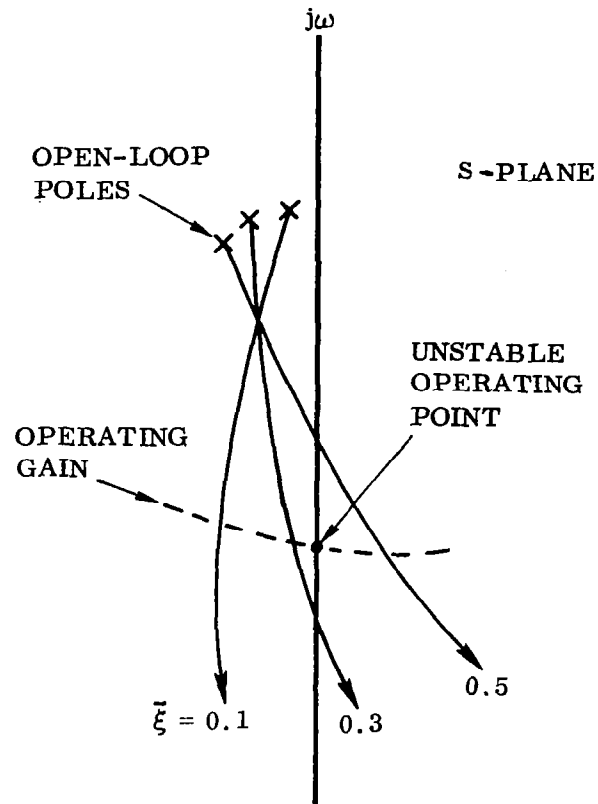


Figure 25. Potentially Unstable Limit Cycle, Root Locus

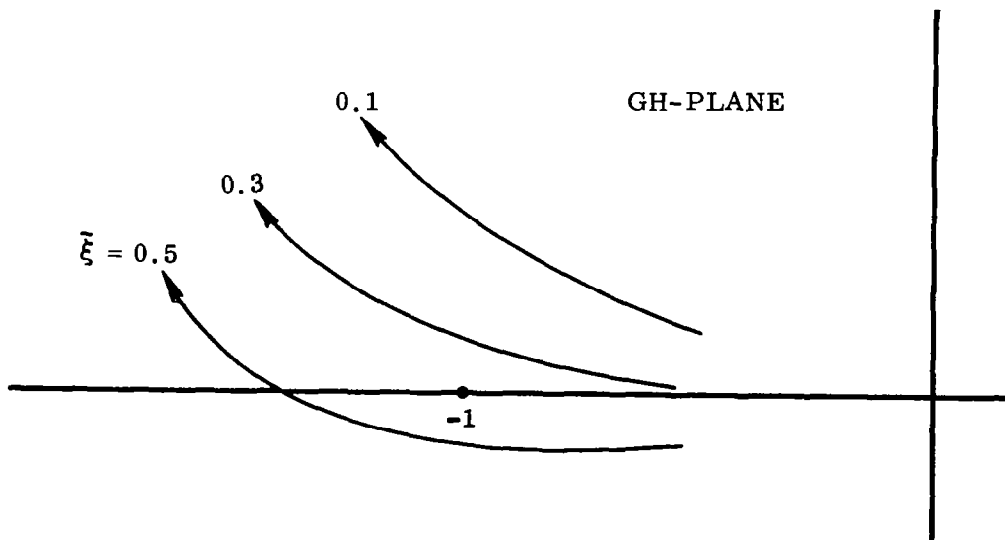


Figure 26. Potentially Unstable Limit Cycle, Nyquist Diagram

A stable limit cycle operating point is shown for an example propellant slosh mode in Figure 27. Here the frequency of the parasitic mode is so low as to make the load torque feedback term negligible; the open-loop pole remains unchanged with changing amplitude of engine motion. The operating point in this case is at $\xi = 0.3$ degree. The same reasoning as before holds. If the amplitude were smaller the root would be unstable, causing divergent oscillations that tend to drive the root toward the $j\omega$ axis. If larger, the oscillations converge, driving the root to the right. The result in either case is an operating point on the $j\omega$ axis.

The restriction on this analysis, that of considering only one frequency at a time, can be mitigated in several ways. One approach would be to compute the open-loop frequency response while re-evaluating the describing function for a particular amplitude for each frequency step. The resulting Nyquist diagram is valid at all frequencies at the cost of considerable computational labor. One can get around this by programming the process for digital computer solution. In principle, the same technique can be applied to solving by means of root locus for roots in the vicinity of the $j\omega$ axis, although the iteration procedure is probably not worth while considering the small increase in accuracy for the increased computations required.

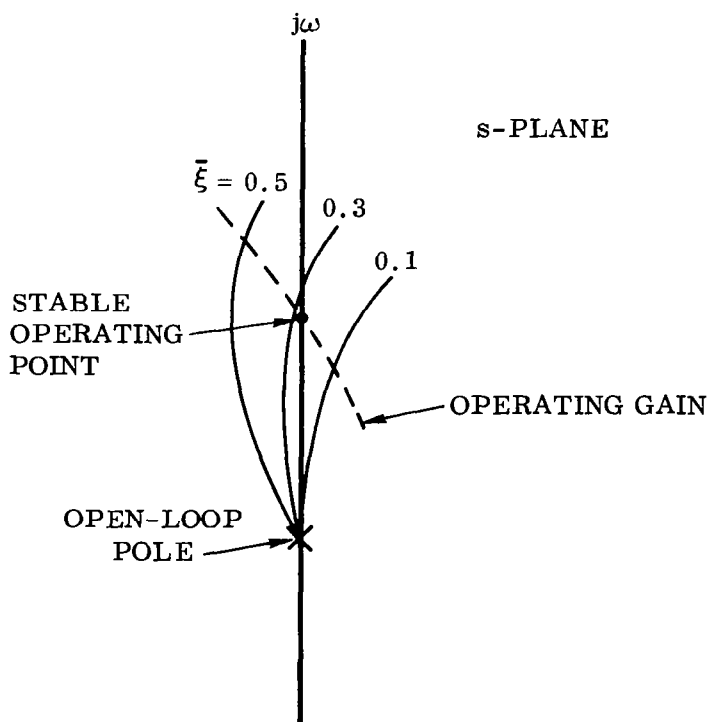


Figure 27. Stable Limit Cycle

A second approach is to approximate the describing function in terms of a single parameter, this parameter then being related to engine amplitude and frequency (Reference 8). Root loci can be plotted for various values of this lag and operating points located in terms of frequency and lag. Referral to plots of K'_C (the lag break frequency) versus $\bar{\xi}$ for various frequencies then determines the amplitude of the possible limit cycle. Unfortunately, this technique is invalid for the higher frequencies, because the approximation becomes progressively poorer.

4.2.7.1.2 Propellant Slosh Damping. Liquid-propelled launch vehicles quite often must rely on antislosh baffles to stabilize these parasitic modes. Paragraph 4.2.6.2 gives a discussion of how these give rise to limit cycles. The interest here is in how the amplitude of these limit cycles may be estimated.

Consider the i^{th} slosh mode loci for various values of K'_C in Figure 28 as a function of an assumed propellant slosh damping, ζ_1 . The point where the dotted line representing the locus of constant operating gain as a function of K'_C crosses the $j\omega$ axis represents a potential limit cycle operating point. The word "potential" is used because it is not known which damping ratio will be effective in the limit cycle. Note that the potential limit cycle frequency drops as the assumed damping increases. This functional relationship is shown in Figure 29. If the damping gets much greater than that shown in the fourth diagram of Figure 28, stable operation is predicted. Since

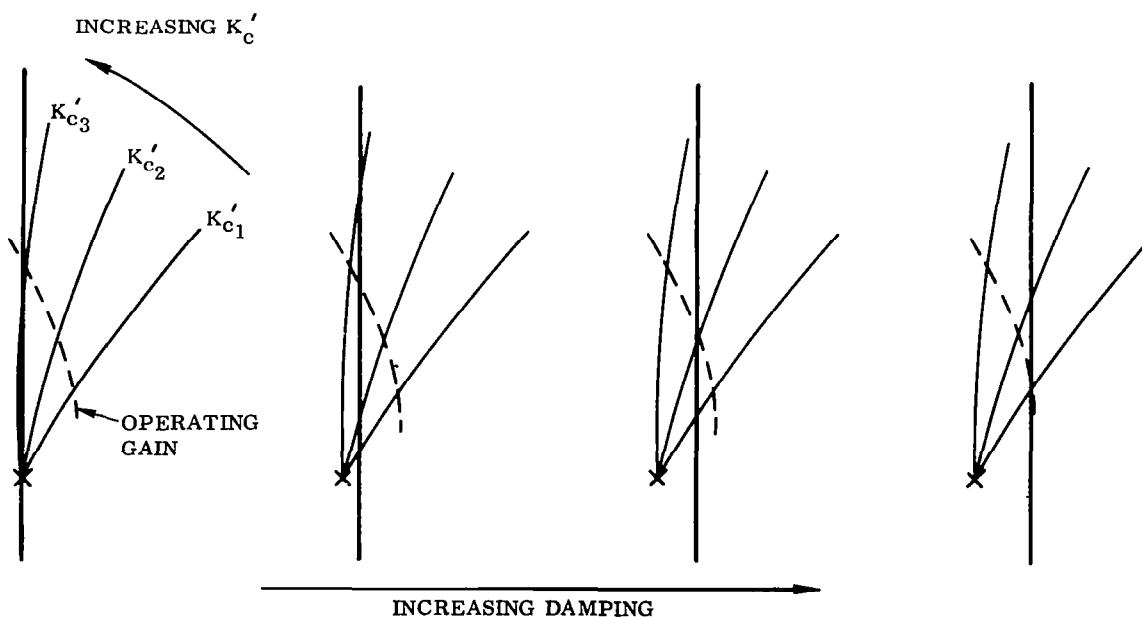


Figure 28. Propellant Slosh Loci Showing Effect of Increasing Slosh Damping

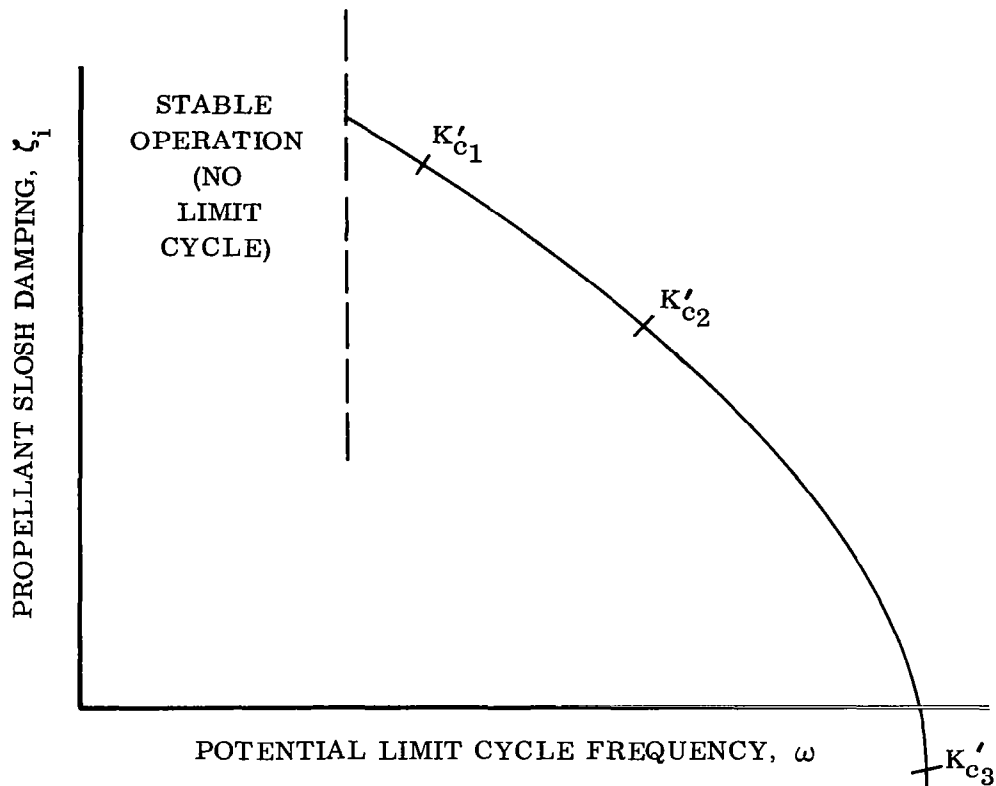


Figure 29. Potential Limit Cycle Frequency Versus Slosh Damping

this means converging oscillations and decreasing damping, this condition cannot occur: the damping must be less than this. Similarly the first diagram in Figure 28, showing zero damping, has an operating point corresponding to a large value of K'_c . This implies large engine amplitudes, hence slosh amplitudes, hence damping greater than zero. One concludes the actual operating condition is at some intermediate point.

Figure 30 shows a typical plot of servo lag frequency K'_c versus engine amplitude $\bar{\xi}$ and frequency ω . Thus for each value of ω and K'_c an engine amplitude can be assigned, and each potential operating point has an associated amplitude of engine motion.

A transfer function relating slosh pendulum amplitude to engine amplitude can be developed for each value of damping assumed. Each potential operating point from Figure 30 represents a particular value of frequency, damping, and engine amplitude. From the transfer function $\frac{\beta_1}{\bar{\xi}}(\zeta_1, s)$, and from these potential operating point data, one can arrive at a particular pendulum amplitude $\bar{\beta}_1$ corresponding to each point. Such a curve might look like that shown in Figure 31. Given the functional relationship between baffle damping, slosh frequency, and slosh amplitude shown in Figure 32, the slosh damping pertaining to each slosh amplitude can be obtained as shown by the dotted line.

At this point we have obtained the slosh damping corresponding to each slosh amplitude. This can be plotted versus potential limit cycle frequency as shown in Figure 33.

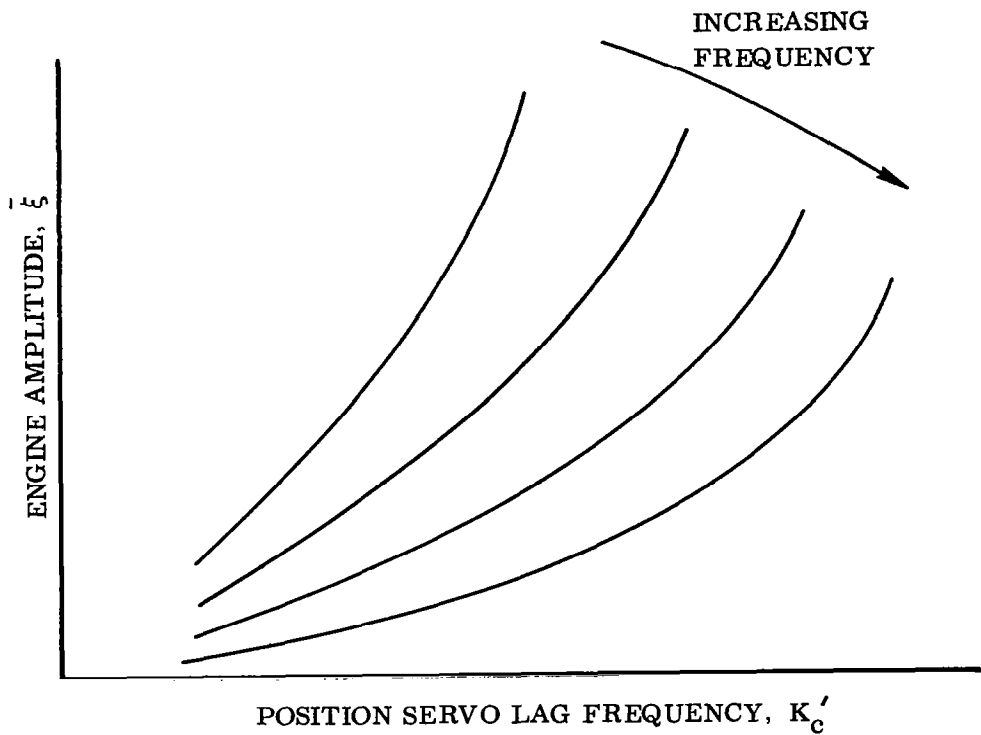


Figure 30. Position Servo Lag Frequency Versus Engine Deflection Half Amplitude

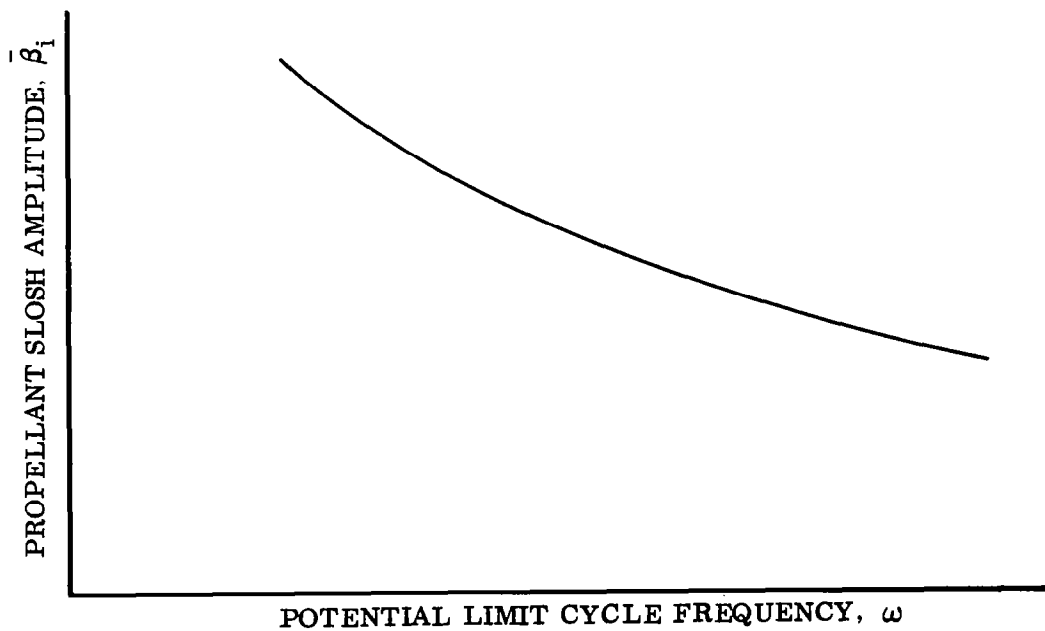


Figure 31. Slosh Amplitude Versus Potential Limit Cycle Frequency

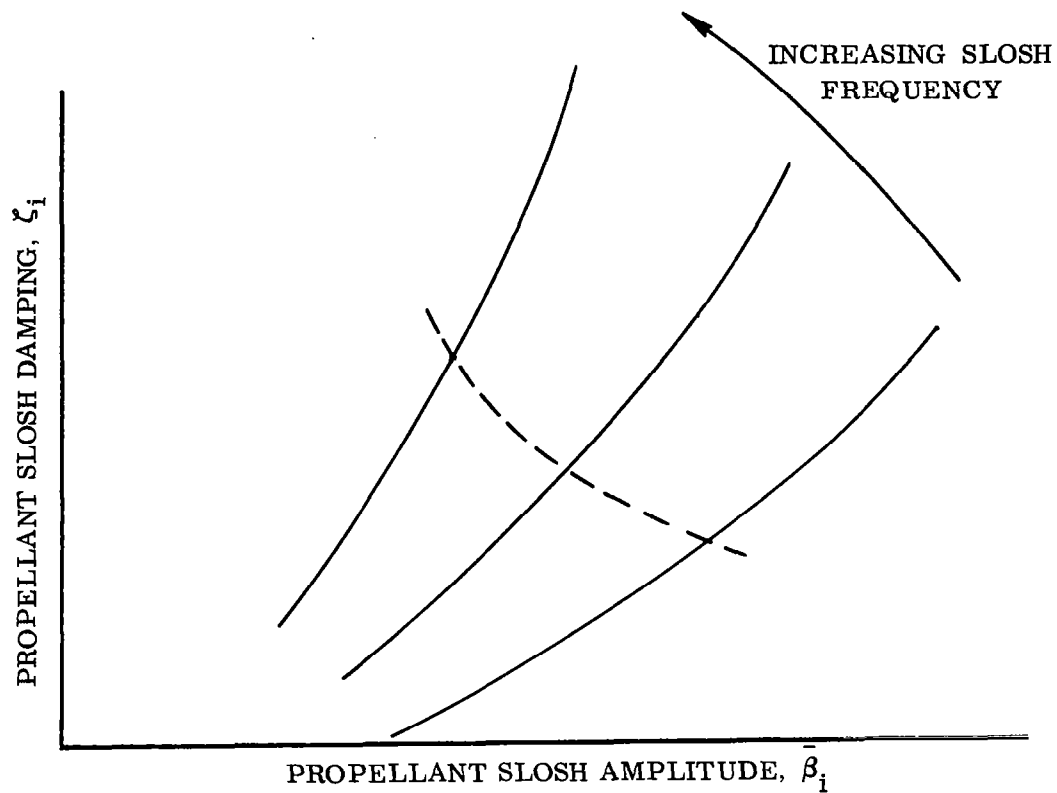


Figure 32. Propellant Slosh Damping Versus Amplitude with Slosh Frequency As a Parameter

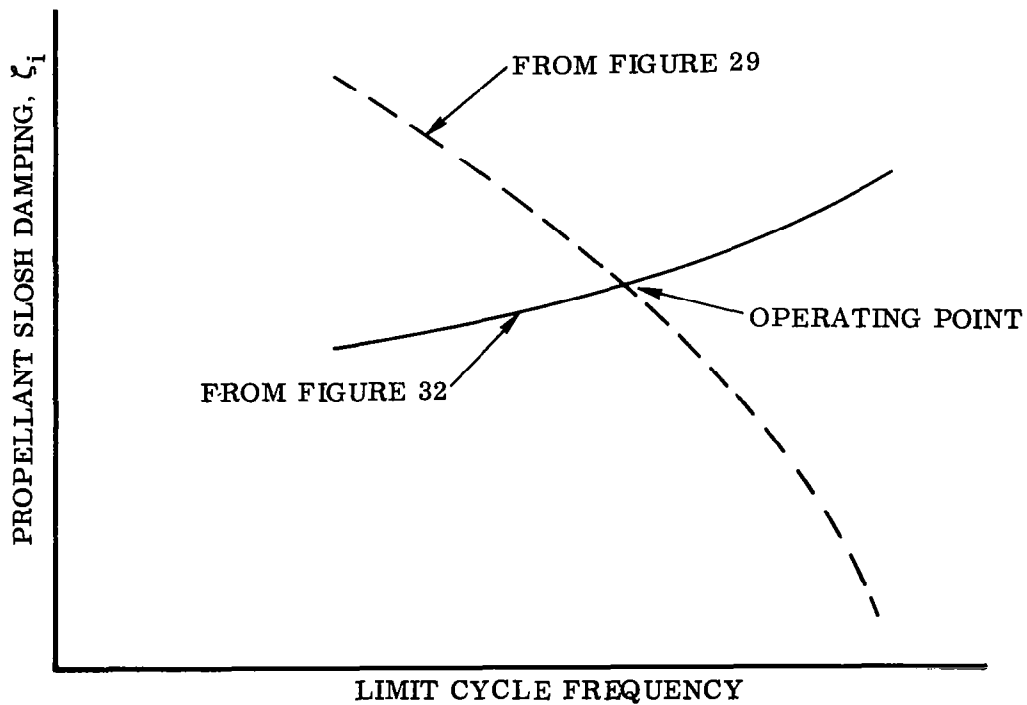


Figure 33. Operating Point for Slosh Limit Cycle

The curve and that of Figure 29 cross at one point shown by the dotted line. This point is the actual limit cycle operating point and corresponds to a particular amplitude and frequency of engine motion.

4.2.7.2 Use of Simulation Techniques. Unfortunately, the analytical technique for predicting slosh limit cycle amplitudes as discussed in Paragraph 4.2.7.1.2 is rather cumbersome and is tedious in application. It can also be inaccurate when compared with flight data. One of the major reasons for this is the time-varying parameter nature of a launch vehicle. For low-frequency limit cycles the rate of convergence or divergence is on the same order as the rate of change of system parameters. The vehicle might pass through a critical flight regime (from the slosh stability standpoint) before the amplitude of slosh motion has had a chance to build up to predicted values. On the other hand, atmospheric disturbances could excite the system to larger amplitudes that die down relatively slowly because of their nearness to the limit cycle condition. A second reason is the approximation made by neglecting load torque feedback on thrust chamber motion. This effect generally increases the gain and phase lag of the loci.

One convenient means for obtaining answers to the questions imposed by the nonlinear behavior of the system is that of direct simulation on an analog or digital computer. Indeed, this is the only way the response of the nonlinear system can be obtained; the describing function technique used here says nothing about system behavior in other than an undisturbed limit cycle condition. The procedure is straightforward; one sets up the equations of motion, solves, and examines the results. On an analog computer the effects of parameter variations can be handled on the spot, and the nonlinearities simulated in full.

4.2.8 GENERAL APPROACH TO PRELIMINARY ANALYSIS AND SYNTHESIS. The preceding paragraphs of this section have been used to introduce various techniques within the current state of the art used in the analysis and synthesis of space launch vehicle flight control systems. The emphasis has been on their application to the preliminary phases of system design; that is, their adaptability to manual calculation. The quality of the basic data available in the early phases of the design process seldom warrants the detailed and more elaborate solutions that rely for the most part on computing machines. This is not to say that machine methods cannot be used at this stage, only that they should not be necessary.

It would be helpful at this point to summarize the general approach used in the initial design phases. Perhaps the most distinguishing feature is the fluidity of the basic data as various vehicle configurations are considered in the light of the launch vehicle mission and the design philosophy used. The control analyst's task in this process is to evaluate the various configurations for their ease of control. Other things being equal the system that is simplest and has the greatest allowable margin in its parameters or vehicle parameters is the most desirable. Considering the preliminary nature of the data, it does not pay to devote lengthy periods of time to exhaustive control analyses.

Rather, an estimate of the flight control system requirements is needed to enable a reasonable judgment of each alternate vehicle configuration.

There are some fixed reference points. For example, the intended range of missions and the design philosophy dictated by these missions are reasonably well established early in the study phase. These will have a strong bearing on the design freedom in configuring the flight control system. If the mission is weapons delivery the highest premium is placed on simplicity, adaptability to mass production techniques, and a minimum of preflight testing. The control system design will reflect this philosophy in its being configured for very wide tolerances in system parameters, with system performance being secondary. The simplest design remaining within the constraints dictated by stability requirements is the probable outcome, with emphasis being on minimum cost per launch. The designer can allow for some percentage of effectivity over a number of launches.

On the other hand, a launch vehicle designed for space missions has a much greater premium placed on maximum performance of all vehicleborne systems, with emphasis being on the greatest probability of mission success for each launch, the total number of launches being small. With the cost per launch being relatively high, one must maximize the reliability with extensive preflight testing and checkout. Each launch is likely to be somewhat different than the last (different payloads, orbital parameters, and the like). Further, the size of this class of launch vehicle is likely to be such that only a near optimum control system design will suffice to give adequate performance. The end result is a rather more complex control system configuration having a narrower operating range and probably incorporating redundancy features.

An additional constraint on the control analyst is the need for allowing sufficient margin in his preliminary definition of flight control system requirements for the changes and refinement in the vehicle design that take place as the configuration crystallizes. He may, for example, be confronted with a marginal situation regarding the necessity for fins that might be brought about by a close proximity of the rigid body control frequency and the first lateral bending mode. It would be better in this situation to indicate the need for fins; it is easier to take them off should they ultimately prove to be unnecessary than it would be to add them later on. In any case, the necessity for those elements of the overall vehicle structure affecting flight control must be indicated early in the game. This situation is an example of a basic dilemma facing the analyst. He cannot provide final system requirements without refined data on vehicle dynamics, and the data are usually not available in detail until the program is well into the ground testing phase. In short, his initial estimates must err on the side of conservatism to avoid, to as great an extent as possible, major structural changes later on.

Emphasis in the preliminary definition of system requirements is on those elements having considerable hardware impact. This includes fins (if they are needed), and if so, what size), thrust vector deflection angle requirements, antislosh baffles (size and

location), bandwidth requirements on the control actuation scheme, sensor locations and characteristics, and autopilot design philosophy (continuously programmed gains and filters, or discrete sequencing of same). The only elements remaining to be specified are those that can be changed relatively easily, such as autopilot gains, limits, and filter requirements. In a computer-controlled system these would be software items.

The recommended procedure in the initial design phase follows. One first estimates the rigid body control requirements in the light of the range of mission requirements. This includes the possibility of auxiliary loops for minimizing drift and/or inflight aerodynamic loads. Derived from these studies will be such things as thrust vector deflection requirements, control frequency, and system gains. Secondly, one examines the lowest-frequency flexible mode in an effort to determine if the low-frequency gain and phase requirements are compatible with the higher-frequency requirements for parasitic modal stability. If these are incompatible one must usually increase the frequency separation between the control and flexible vehicle modes; stiffen the structure or add fins. If the latter, then the rigid body control stability must be re-evaluated at the lower gains that result. The autopilot sensor requirements and control actuation bandwidth are found in this portion of the effort. Thirdly, one investigates (in liquid-propelled vehicles) the propellant sloshing stability to define baffling requirements and possible modifications to the flight control system. The end result of these studies provides an initial specification of the flight control system requirements. The most critical of these is the requirements for stability of the parasitic modes. As better data become available on the modal parameters, investigation of the effects of changes in these data take first priority as they may influence the final choice of sensor characteristics and locations, autopilot gains, and filter parameters.

4.3 ANALYSIS AND SYNTHESIS - DETAILED DESIGN PHASE

The preceding subsection has outlined the techniques used in roughing out major features of the launch vehicle flight control system. In the detailed design and analysis phase, the techniques used are basically the same, differing only in the detail and comprehensiveness of the studies. Hand calculations no longer suffice because of the number of degrees of freedom that are taken into account simultaneously; one must rely on machine computation to obtain solutions to the desired degree of accuracy.

One of the greatest problems facing the control engineer is that of determining which details in the mathematical description are significant for a particular analysis, and which can be ignored. With this in mind, the first portion of this subsection deals with the equations of motion for the launch vehicle and the approximations that can be made to simplify the set of simultaneous differential equations. It is not, however, the intent of this discussion to provide a detailed derivation of a universally applicable set of equations. Such derivations are covered in other monographs in this series. Instead, it is oriented toward enabling the control engineer to derive the equations pertinent to the particular vehicle he is analyzing. This is done via a demonstration

of such a derivation along with a discussion of the approximations made. The vehicle is a medium sized, liquid-propelled launch vehicle that obtains its control from gimbaling the thrust chambers to obtain thrust vector deflection. This is representative of the more complex dynamics that may be encountered. The second portion deals with ways and means used to obtain solutions to the detailed equations of motion; that is, methods of detailed analysis and synthesis of the flight control system taking into account all pertinent parameters.

4.3.1 DEVELOPMENT OF EQUATIONS OF MOTION. For stability studies, the dependent variables are perturbational quantities. The measured variables consist of two parts, a steady-state component that varies slowly as a function of time and a perturbational component that varies relatively rapidly at frequencies corresponding to the various modes of oscillation of the launch vehicle. The former quantities are the subject of interest in trajectory studies, the latter in dynamic stability studies. Coupling between the two is small for the large, relatively slow moving launch vehicles under discussion and is usually neglected, thus decoupling the short-period perturbational equations from the long-period equations of the trajectory-dependent terms. The two are generally combined only in comprehensive simulations of a launch vehicle flight through the atmosphere where flight loads are being studied.

The equations of motion developed in this subsection are not intended to be universally applicable to any flexible space launch vehicle. Rather, they are intended to be representative of the kind of equations that might be developed to suit the requirements of the particular case. In this discussion a liquid-propelled launch vehicle with gimballed rockets engines is assumed. This is typical of the more complex situations. A simpler example would be a solid propellant launch vehicle using secondary fluid injection for thrust vector deflection.

4.3.1.1 Basic Principles. The large launch vehicle is an elastic structure that will deform under the action of externally applied loads. This sets up an elastic strain energy within the structure that tends to restore the vehicle to its undeflected shape. This cannot occur instantaneously as the restoring motion is resisted by the inertial loading of the structure itself. In effect, and in fact, the vehicle is a continuous structure of distributed inertial and elastic properties. In addition there are sloshing propellants and certain discrete spring-mass systems (engines, turbopumps, and the like) that contribute to the overall dynamic characteristics. The success of the control analyst in configuring a flight control system is heavily dependent on the mathematical description of the vehicle. Such a description always involves some degree of approximation. Which approximations to use becomes a major portion of the analysis and synthesis task.

The usual launch vehicle has a high degree of structural (hence elastic and inertial) symmetry. The control and inertial axes nearly coincide with the geometric axes of the structure. Thus inertial and elastic coupling effects between the control axes are small. Similarly, the aerodynamic properties are such as to result in negligible cross-coupling effects between planes. In short, coupling between control planes is very small, so small as to permit almost exclusive reliance on planar analyses of vehicle stability. In this monograph planar analysis will be assumed. Clearly, if the particular configuration requires it this approach would have to be modified.

The equations of motion are usually based on the Rayleigh-Ritz method (Reference 9) of employing assumed structural mode shapes. Since the launch vehicle is a continuous structure (at least in part) it has an infinite number of degrees of freedom. To reduce the problem to manageable proportions, a modal solution approach is indicated; the vehicle deflected shape under forced vibration (due to action of the aerodynamic environment and the flight control system) is approximated by a linear combination of a finite number of assumed mode shapes. A premium is placed on proper selection of these mode shapes. The better the selection, the smaller is the number of degrees of freedom necessary to adequately represent the dynamic behavior of the vehicle.

Solution of the hydrodynamic equations describing fluid motion within a rigid walled tank undergoing both translation and rotation provides a general solution for the sloshing forces in terms of a superposition of sloshing modes. For each slosh mode there is a particular characteristic frequency and pressure distribution on the walls of the tank. The integral of this pressure distribution results in the slosh moments and lateral slosh forces acting on the tank. In addition, the inertial properties of the tank and fluid contents are modified. The form of the equations of motion is such that the slosh forces and moments can be represented by a mechanical analogy. The two most popularly used are the spring-mass analogy and the pendulum analogy. In this development the former is used, although the latter is more convenient when analyzing sloshing stability alone as in Paragraph 4.2.6.

Generally speaking, only the first lateral slosh mode is considered. It has the lowest frequency, the least inherent fluid damping, and the greatest effective mass of moving fluid; hence the greatest effect on the vehicle. However, certain situations may call for considerations of other modes, for example:

1. The second lateral slosh mode may be significant relative to the first if the tank geometry is conducive to second mode motion. Internal structures such as a hemispherically domed bottom in a cylindrical tank can have this effect at low fluid levels.
2. A segmented tank generally has different slosh frequencies, depending on the direction of excitation (the control plane being considered).

There is no general rule. The propellant slosh parameters of mass, spring constant, and moment coefficient together with the modifications to the overall vehicle inertial properties (the so-called "reduced" mass, center of gravity, and moments of inertia) for the first few slosh modes should be evaluated for their potential influence on the problem. In particular, the mass associated with the second and higher slosh modes for right circular cylindrical tanks is only a few percent at the first mode for all fluid levels except the near empty condition, and therefore is neglected. Consideration of the first fluid slosh mode is sufficient for most vehicle geometries at most flight times.

Slosh motion in only one plane is considered. This is made possible by the cylindrical symmetry of most space launch vehicles; the center of gravity is quite close to the vehicle centerline and the cross products of inertia are small, resulting in the pitch, yaw, and roll control planes being largely uncoupled. In cases where this is not so the system should be examined for the possibility of exciting swirl modes: rotation of the fluid about the tank longitudinal axis. Proper design should not permit such a circumstance; the forces generated can become very large and the damping is quite low. Experience to date on the present range of launch vehicles would indicate that swirl is usually of no concern.

The forcing functions acting on the launch vehicle consist of the aerodynamic and propulsive forces as well as the control actuation torques (significant when gimballed thrust chambers are used for control). The propellant slosh forces also may be considered as "external" forces, or else incorporated in the determination of the mode shapes used in the analysis.

The launch vehicle will also contain dissipative forces leading to some damping of the elastic and sloshing degrees of freedom (the damping in the gimbal bearing is taken into account separately in the equation of engine motion). The dissipative energy is very small compared to the kinetic and potential energies for the lower-frequency elastic modes, and consequently has negligible effect on these modes. A small viscous damping term is added to the equations of motion to account for any energy dissipation thought to be significant. For a more complete discussion of the dynamic modeling of the launch vehicle modes and the influence of damping terms, see Reference 8.

4.3.1.2 Example - Equations of Motion for Medium Sized, Liquid Propelled Launch Vehicle. The following derivation of the equations of motion is restricted to those launch vehicles enjoying a relatively wide separation between rigid body and propellant slosh frequencies and the flexible vehicle modes. The undeflected vehicle axis system is treated as an inertial reference frame for describing slosh and flexible vehicle motion. The rigid body, propellant slosh, and flexible vehicle modes are thus all artificially uncoupled, causing inertial and elastic crosscoupling terms between these

various degrees of freedom. The assumed mode shapes for the vehicle are derived from a lumped parameter model that includes the gimbaled engine masses. Thus inertial and elastic crosscoupling terms appear between the various flexible vehicle modes. Slosh masses are not included in the lumped parameter model.

Considering first the flexible body degrees of freedom, Lagrange's equation for the k^{th} flexible mode excluding the gimbaled engines is

$$\frac{d}{dt} \left(\frac{\partial T_V}{\partial \dot{q}^{(k)}} \right) - \frac{\partial T_V}{\partial q^{(k)}} + \frac{\partial U_V}{\partial q^{(k)}} = Q^{(k)} = \frac{\partial W_e}{\partial q^{(k)}} \quad (164)$$

where

$T_V = T - T_E$, the kinetic energy of the vehicle excluding the gimbaled engines

$U_V = U - U_E$, the potential energy of the vehicle excluding the gimbaled engines

$Q^{(k)}$ = generalized force exciting the k^{th} mode

W_e = virtual work done by external forces under a virtual deflection of the generalized coordinate, $q^{(k)}$

For perturbational motion within the linear range, the rate of change of the kinetic energy with modal displacement is zero. The kinetic and potential energies of the launch vehicle including engines are given by

$$T = \frac{1}{2} \sum_{i=1}^n m^{(i)} [\dot{q}^{(i)}]^2 \quad (165)$$

$$U = \frac{1}{2} \sum_{i=1}^n m^{(i)} [\omega^{(i)} q^{(i)}]^2 \quad (166)$$

where $m^{(i)}$ and $\omega^{(i)}$ are the generalized mass and frequency of the i^{th} orthogonal normal mode of the vehicle. No elastic or inertial crosscoupling terms appear because of the orthogonality property. Equation 164 may be written as

$$m^{(k)} \ddot{q}^{(k)} - \frac{d}{dt} \left(\frac{\partial T_E}{\partial \dot{q}^{(k)}} \right) + m^{(k)} [\omega^{(k)}]^2 q^{(k)} - \frac{\partial U_E}{\partial q^{(k)}} = Q^{(k)} \quad (167)$$

In obtaining expressions for T_E and V_E , assume one equivalent gimbaled engine (the gimbaled engines of a multiple-engine launch vehicle are usually identical) and one equivalent fixed engine.

$$T_E = \frac{1}{2} M_R \left\{ \sum_{i=1}^n \left[\phi_{xT}^{(i)} - \ell_R \sigma_{xE}^{(i)} \right] \dot{q}^{(i)} \right\}^2 + \frac{1}{2} I_{R0} \left\{ \sum_{i=1}^n \sigma_E^{(i)} \dot{q}^{(i)} \right\}^2 \quad (168)$$

$$U_E = \frac{1}{2} K_\alpha \left\{ \sum_{i=1}^n \beta_{xT}^{(i)} q^{(i)} \right\}^2 \quad (169)$$

where

$$\sum_{i=1}^n \beta_{xT}^{(i)} q^{(i)} = \sum_{i=1}^n \left(\sigma_{xT}^{(i)} - \sigma_{xE}^{(i)} \right) q^{(i)} \quad (170)$$

is the engine deflection relative to the aft end of the vehicle (elastic axis at the gimbal point for the orthogonal normal modes of the launch vehicle including engines; see Figure 34. The torsional spring constant K_α is a combination of actuator, actuator support structure, and hydraulic spring constants. It is also the torsional spring used in the lumped parameter model from which the orthogonal normal mode shapes $\phi_x^{(i)}$

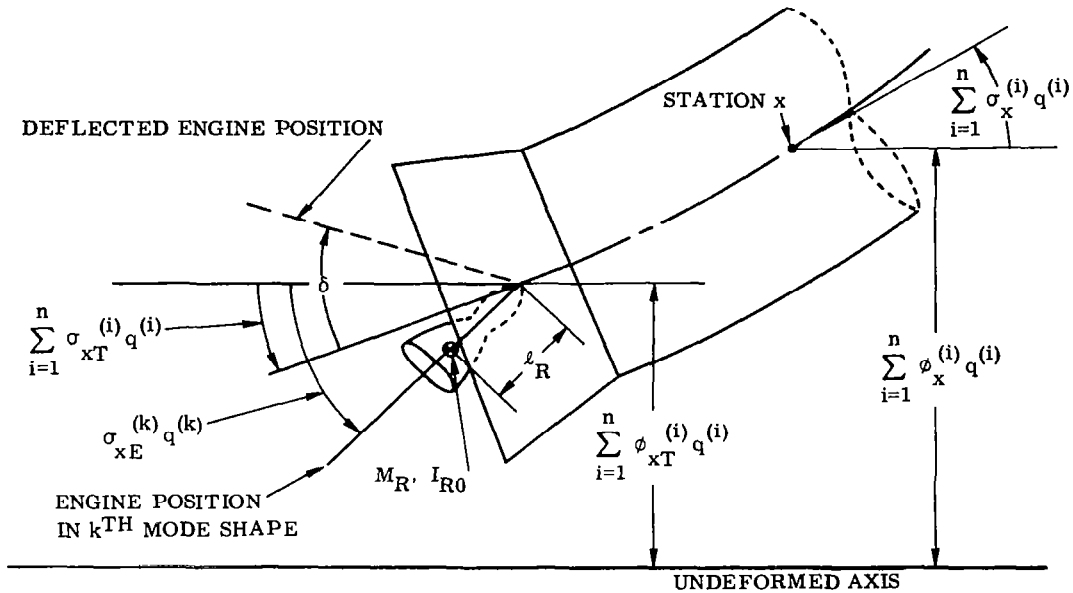


Figure 34. Coordinate System at Aft End of Launch Vehicle for Modal Deflections, Pitch Plane

are obtained. In effect, if the engine is commanded to null, K_α represents the influence coefficient relating thrust chamber deflection from null to an externally applied torque about the gimbal axis. It is, in general, a frequency-dependent variable. Performing the indicated operations of Equation 167, the following expressions are obtained.

$$\frac{d}{dt} \left(\frac{\partial U_E}{\partial \dot{q}^{(k)}} \right) = \sum_{i=1}^n \left\{ M_R \left[\phi_{xT}^{(k)} + \ell_R (\beta_{xT}^{(k)} - \sigma_{xT}^{(k)}) \right] \phi_{xT}^{(i)} + \left[M_R \ell_R \phi_{xT}^{(k)} + I_R (\beta_{xT}^{(k)} - \sigma_{xT}^{(k)}) \right] (\beta_{xT}^{(i)} - \sigma_{xT}^{(i)}) \right\} \ddot{q}^{(i)} \quad (171)$$

$$\frac{\partial U_E}{\partial q^{(k)}} = K_\alpha \beta_{xT}^{(k)} \sum_{i=1}^n \beta_{xT}^{(i)} q^{(i)} \quad (172)$$

where $I_R = I_{R0} + M_R \ell_R^2$ is the rocket engine moment of inertia about the gimbal axis.

The generalized force acting on each mode is composed of several components as given below:

1. **Modal Damping** - The internal damping of each structural mode arises from material strain hysteresis, Coulomb friction in structural joints, and viscous propellant action within the tanks. These forces are approximated by an equivalent small viscous friction term added to each structural mode, with no crosscoupling between the modes assumed. Thus

$$Q_{(\text{damping})}^{(k)} = -2\zeta^{(k)} \omega^{(k)} m_l^{(k)} \dot{q}^{(k)}$$

This simplified treatment is justified by the fact that the actual damping is low (ranging from a fraction of a percent of critical to about two percent). The damping is found by test to produce little coupling between the normal modes of the structure. In addition, it can be shown (Reference 8) that the structural damping leads to no crosscoupling in the equations of motion.

2. **Aerodynamics** - In writing the generalized force due to aerodynamics, a Procrustean view is adopted in that quasi-steady aerodynamics are assumed. The mathematical model assumes the forces to be directly proportional to and in phase with the local angle of attack at any point on the launch vehicle. Thus

$$Q_{(\text{aero})}^{(k)} = (qS) \sum_x C_{(N/\alpha)_x} \phi_x^{(k)} \alpha_x \quad (173)$$

where the local angle of attack is

$$\alpha_x = \alpha + \sum_{i=1}^n \left[\sigma_x^{(i)} q^{(i)} - \frac{1}{V} \phi_x^{(i)} \dot{q}^{(i)} \right] - \frac{x - x_{cg}}{V} \dot{\theta} \quad (174)$$

and the rigid body angle of attack is given by

$$\dot{\alpha} = \dot{\alpha}_W + \dot{\theta} + \frac{1}{V} [\dot{z} - A_x \theta] \quad (175)$$

The q in Equation 173 is dynamic pressure; S is the reference area. Since q carries no superscript, there should be no confusion with the generalized modal coordinate $q^{(i)}$. It has been assumed that there are no aerodynamic forces acting on the engine itself. The engines are gimballed within a shroud or skirt that shields them from the airstream.

The use of quasi-steady aerodynamic theory for the low-frequency rigid body motions of the vehicle is correct when steady-state experimental or empirical distributions of $C_{(N/\alpha)_x}$ are used. At the higher frequencies associated with

the flexible modes of the airframe this theory is justified only in that there exists no suitable alternative. Fortunately, for launch vehicles having no aerodynamic surfaces the aerodynamic effects on the higher-frequency bending modes are rather small. The consequences of this assumption have a relatively small effect from the overall point of view. Note that the engines are assumed to be shrouded in such a way that there are no aerodynamic forces acting on the thrust chambers.

3. Gimbaled Engine Dynamics - The motion of the gimbaled thrust chamber produces inertial reaction forces and torques on the launch vehicle due to translational and rotational accelerations of the engine mass. Thus the lateral force at the gimbal axis due to lateral acceleration of the engine center of mass is

$$\begin{aligned} F_{(\text{engine})} = & M_R [\ddot{z} + (\ell_c + \ell_R) \ddot{\theta} - \ell_R \ddot{\delta}] \\ & - M_R \sum_{i=1}^n \left[\phi_{xT}^{(i)} - \ell_R \sigma_{xT}^{(i)} \right] \ddot{q}^{(i)} \end{aligned} \quad (176)$$

The torque acting on the vehicle due to positive rotational acceleration of the engine is

$$\begin{aligned} m_{(\text{engine})} = & I_R \ddot{\delta} - (I_R + M_R \ell_R \ell_c) \ddot{\theta} - M_R \ell_R \ddot{z} \\ & + \sum_{i=1}^n \left[M_R \ell_R \phi_{xT}^{(i)} - I_R \sigma_{xT}^{(i)} \right] \ddot{q}^{(i)} \end{aligned} \quad (177)$$

This torque is provided by the actuator load pressure and resisted by gimbal friction.

$$m_{(\text{engine})} = A R P_L - C_V \dot{\delta} - C_B \frac{\dot{\delta}}{|\dot{\delta}|} \quad (178)$$

The load torque T_L acting to resist engine motion is therefore (from Equations 177 and 178)

$$\begin{aligned} T_L = & - (I_R + M_R \ell_R \ell_c) \ddot{\theta} - M_R \ell_R \ddot{z} \\ & + \sum_{i=1}^n \left[M_R \ell_R \phi_{xT}^{(i)} - I_R \sigma_{xT}^{(i)} \right] \ddot{q}^{(i)} \end{aligned} \quad (179)$$

Returning to the present discussion, the virtual work due to a deflection $\Delta q^{(k)}$ is

$$\begin{aligned}
\Delta W_{(\text{engine})} = & \left[I_R \ddot{\theta} - (I_R + M_R \ell_R \ell_c) \ddot{\theta} - M_R \ell_R \ddot{z} \right] \sigma_{xT}^{(k)} \Delta q^{(k)} \\
& + M_R \left[\ddot{z} + (\ell_c + \ell_R) \ddot{\theta} - \ell_R \ddot{\theta} \right] \phi_{xT}^{(k)} \Delta q^{(k)} \\
& + \sum_{i=1}^n \left[M_R \ell_R \phi_{xT}^{(i)} - I_R \sigma_{xT}^{(i)} \right] \ddot{q}^{(i)} \sigma_{xT}^{(k)} \Delta q^{(k)} \\
& - \sum_{i=1}^n \left[M_R (\phi_{xT}^{(i)} - \ell_R \sigma_{xT}^{(i)}) \right] \ddot{q}^{(i)} \phi_{xT}^{(k)} \Delta q^{(k)} \quad (180)
\end{aligned}$$

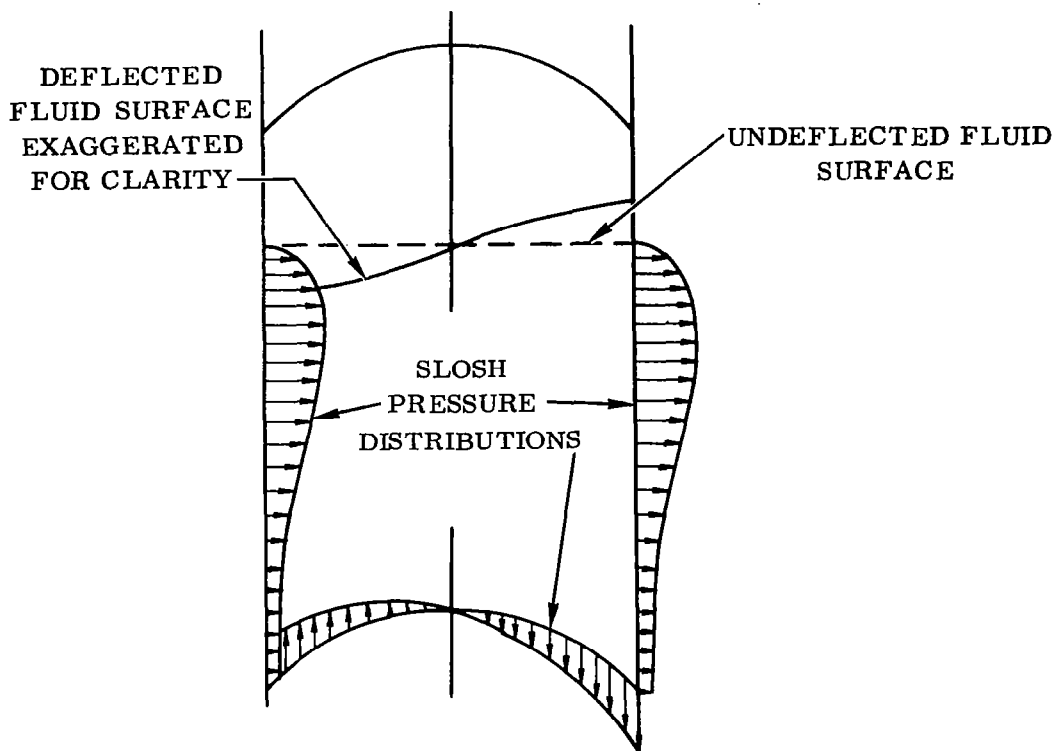
Thus the generalized force is

$$\begin{aligned}
Q_{(\text{engine})}^{(k)} = & \left[I_R \sigma_{xT}^{(k)} - M_R \ell_R \phi_{xT}^{(k)} \right] \ddot{\theta} + M_R \left[\phi_{xT}^{(k)} - \ell_R \sigma_{xT}^{(k)} \right] \ddot{z} \\
& + \left[M_R (\ell_R + \ell_c) \phi_{xT}^{(k)} - (I_R + M_R \ell_R \ell_c) \sigma_{xT}^{(k)} \right] \ddot{\theta} \\
& - \sum_{i=1}^n \left[M_R (\phi_{xT}^{(k)} - \ell_R \sigma_{xT}^{(k)}) \phi_{xT}^{(i)} + (I_R \sigma_{xT}^{(k)} \right. \\
& \left. - M_R \ell_R \phi_{xT}^{(k)}) \sigma_{xT}^{(i)} \right] \ddot{q}^{(i)} \quad (181)
\end{aligned}$$

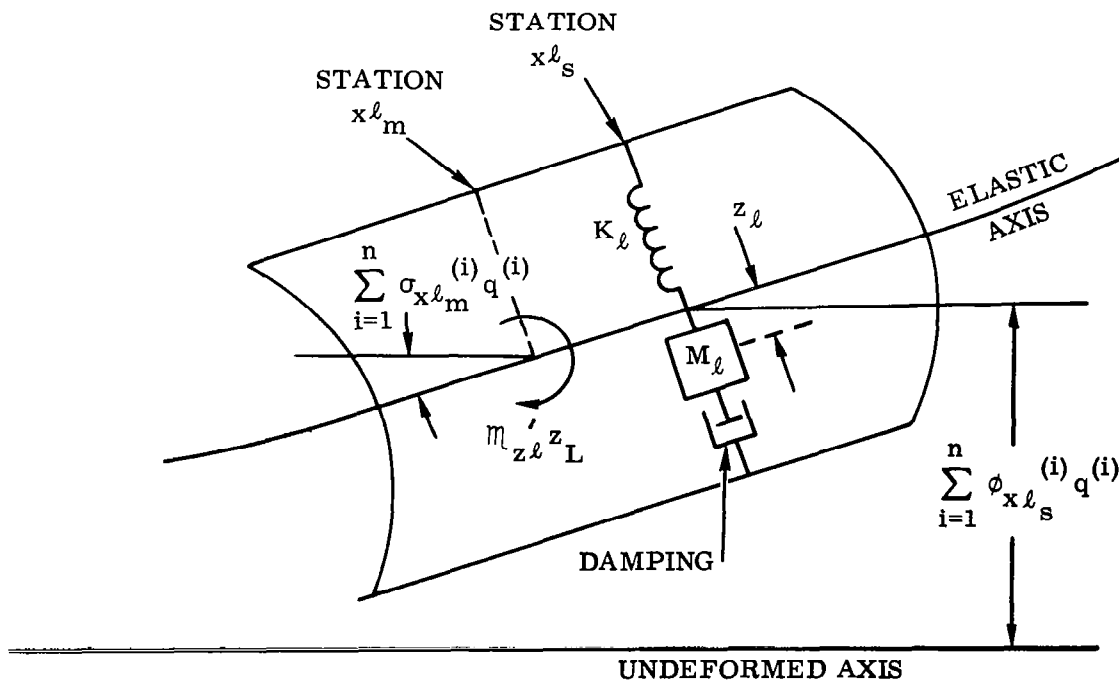
Note that the effect of axial acceleration on the engine mass has been neglected.

4. Slosh Forces - Figure 35 illustrates the coordinate system pertaining to the ℓ^{th} propellant tank. Only the first mode of fluid slosh is assumed to be effective in producing significant forces on the vehicle structure. This restriction can be removed by letting the total number of slosh tanks, m , be the total number of slosh modes considered in all the tanks.

The mathematical model consists of a spring-mass system with damping located at station $x\ell_s$, and an applied moment proportional to the deflection of the slosh mass at station $x\ell_m$. The spring-mass is located at the center of lateral slosh pressure on the walls of the tank. The moment is of such a size and location to correspond to the sloshing moment produced on the structure. These are approximations of the real, distributed forces and moments produced by perturbational (small amplitude) sloshing that would be obtained by integrating the slosh pressure distributions over the sides and bottom of the propellant tank. The spring-mass-couple parameters can be related to the pendulum parameters used in Paragraph 4.2.6, although not uniquely. The choice of spring-mass location (a distance $\ell_{\ell s}$ forward of the "reduced" center



a. First-Mode Slosh Pressure Distribution, ℓ^{th} Tank



b. Spring-Mass-Couple Model, ℓ^{th} Tank

Figure 35. Coordinate System at ℓ^{th} Propellant Tank, Spring-Mass-Couple Analogy .

of gravity) and the location of the applied moment (a distance ℓ_{lm} forward of the "reduced" center of gravity) depend upon the slosh pressure distribution within the tank. Given ℓ_{ls} , the two moment coefficients m'_{zl} and m'_{ql} are determined uniquely (Table 8). The subscript ℓ has replaced the subscript i used in Paragraph 4.2.6.

Table 8. Slosh Parameter Equivalents

DEFINITION	EQUIVALENT
Mass, M_ℓ	M_ℓ
Frequency, ω_ℓ	ω_ℓ
Spring Constant, K_ℓ	$\omega_\ell^2 M_\ell$
Damping Ratio, ζ_ℓ	ζ_ℓ
Moment Coefficient, m'_{zl}	$(\ell_\ell - \ell_{ls}) K_\ell$
Moment Coefficient, m'_{ql}	$(\ell_\ell - \ell_{ls} - L_\ell) M_\ell$

Note that a simpler, although more inaccurate, analogy results when m'_{ql} is set to zero by locating the spring mass at $\ell_{ls} = (\ell_\ell - L_\ell)$.

The virtual work, neglecting the effect of the slosh damping force on the vehicle, is given by

$$\Delta W_{(slosh)}^{(k)} = - \sum_{\ell=1}^m K_{\ell} z_{\ell} \phi_{xs}^{(k)} \Delta q^{(k)} - \sum_{\ell=1}^m m'_{zl} z_{\ell} \sigma_{xm}^{(k)} \Delta q^{(k)} \quad (182)$$

The generalized force on the k^{th} flexible vehicle mode is given by

$$Q_{(slosh)}^{(k)} = - \sum_{\ell=1}^m \left[K_{\ell} \phi_{xs}^{(k)} + m'_{zl} \sigma_{xm}^{(k)} \right] z_{\ell} \quad (183)$$

The equation of motion of the ℓ^{th} sloshing mass is given by

$$M_{\ell} \left[\ddot{z}_{\ell} + 2\zeta_{\ell} \omega_{\ell} \dot{z}_{\ell} + \omega_{\ell}^2 z_{\ell} \right] = - M_{\ell} \ddot{z} + \left[M_{\ell} \ell_{ls} + m'_{ql} \right] \ddot{\theta} + \sum_{i=1}^n \left[M_{\ell} \phi_{xs}^{(i)} + m'_{ql} \sigma_{xm}^{(i)} \right] \ddot{q}^{(i)} \quad (184)$$

where $\omega_{\ell}^2 = K_{\ell}/M_{\ell}$. K_{ℓ} and the remaining coefficients M_{ℓ} , ℓ_{ℓ} , m'_{ql} , m'_{zl} are defined in terms of the pendulum analogy for propellant sloshing in Paragraph 4.2.6. It is sufficient to note here the crosscoupling between rigid body,

propellant slosh, and flexible vehicle bending evident from Equations 183 and 184.

5. **Propulsive Forces** - The orthogonal normal modes determined from the lumped parameter model represent conditions where there is no net work being done on the structure. Thus an engine deflection with respect to the undeformed axis of the vehicle of $\sigma_{xE}^{(k)} q^{(k)}$ will do no work on the k^{th} flexible mode. Any other engine angle will either take energy out of, or put energy into, the mode. Thus the generalized force acting on the k^{th} bending mode is given by

$$\begin{aligned} Q_{(thrust)}^{(k)} = & -T_c \left[\delta - \sum_{i=1}^n \sigma_{xT}^{(i)} q^{(i)} + \sigma_{xE}^{(k)} q^{(k)} \right] \phi_{xT}^{(k)} \\ & + T_f \left[\sum_{i=1}^n \sigma_{xT_f}^{(i)} q^{(i)} - \sigma_{xE_f}^{(k)} q^{(k)} \right] \phi_{xT_f}^{(k)} \end{aligned} \quad (185)$$

This is an approximation to the actual situation where more terms would have to be included to account for virtual work done in both the lateral and longitudinal directions.

The equation of motion for the k^{th} elastic mode can now be written:

$$\begin{aligned} & \left[\mathfrak{M}^{(k)} - M_R \ell_R \theta_{xT}^{(k)} \phi_{xT}^{(k)} - \left(M_R \ell_K \phi_{xT}^{(k)} - I_R \sigma_{xE}^{(k)} \right) \beta_{xT}^{(k)} + I_R \beta_{xT}^{(k)} \sigma_{xT}^{(k)} \right] \ddot{q}^{(k)} \\ & + \left[2 \zeta^{(k)} \omega^{(k)} \mathfrak{M}^{(k)} + \frac{qS}{V} \sum_x \left(C_{(N/\alpha)_x} \phi_x^{(k)} \phi_x^{(k)} \right) \right] \dot{q}^{(k)} + \left[\mathfrak{M}^{(k)} (\omega^{(k)})^2 \right. \\ & \left. - T_c \beta_{xT}^{(k)} \phi_{xT}^{(k)} - T_f \beta_{xT_f}^{(k)} \phi_{xT_f}^{(k)} - (qS) \sum_x \left(C_{(N/\alpha)_x} \phi_x^{(k)} \sigma_x^{(k)} \right) \right] q^{(k)} \\ = & \left(I_R \sigma_{xT}^{(k)} - M_R \ell_R \phi_{xT}^{(k)} \right) \ddot{\delta} - T_c \phi_{xT}^{(k)} \delta + M_R \left[\phi_{xT}^{(k)} - \ell_R \sigma_{xT}^{(k)} \right] \ddot{z} \\ & + \left[M_x (\ell_R + \ell_c) \phi_{xT}^{(k)} - (I_R + M_R \ell_R \ell_c) \sigma_{xT}^{(k)} \right] \ddot{\theta} \\ & + (qS) \sum_x \left[C_{(N/\alpha)_x} \phi_x^{(k)} \left(\alpha - \frac{x-x_{cg}}{V} \dot{\theta} \right) \right] + \sum_{\substack{i=1 \\ i \neq k}}^n \left[M_R \ell_R \beta_{xT}^{(k)} \phi_{xT}^{(i)} \right. \\ & \left. + \left(M_R \ell_R \phi_{xT}^{(k)} - I_R \sigma_{xE}^{(k)} \right) \beta_{xT}^{(i)} - I_R \beta_{xT}^{(k)} \sigma_{xT}^{(i)} \right] \ddot{q}^{(i)} \end{aligned}$$

$$\begin{aligned}
& - \sum_{i=1}^n \frac{qS}{V} \sum_x \left[C_{(N/\alpha)_x} \phi_x^{(k)} \phi_x^{(i)} \right] \dot{q}^{(i)} \\
& \quad i \neq k \\
& + \sum_{i=1}^n \left[(qS) \sum_x \left(C_{(N/\alpha)_x} \phi_x^{(k)} \sigma_x^{(i)} \right) + T_c \phi_{xT}^{(k)} \sigma_{xT}^{(i)} + T_f \phi_{xT_f}^{(k)} \sigma_{xT_f}^{(i)} \right] q^{(i)} \\
& \quad i \neq k \\
& - \sum_{\ell=1}^m \left[K_{\ell} \phi_{x\ell}^{(k)} + m'_{z\ell} \sigma_{x\ell} \right] z_{\ell} \quad (186)
\end{aligned}$$

The equations of motion for the rigid body degrees of freedom can be developed in the same manner, the only difference being in the generalized forces due to thrust. Thus for rigid body pitch,

$$\begin{aligned}
I_0 \ddot{\theta} + m_{\theta} \dot{\theta} - F_{\alpha} \ell_{\alpha} \alpha &= (I_R + M_R \ell_R \ell_c) \ddot{\delta} + T_c \ell_c \delta \\
& - \sum_{i=1}^n (I_R + M_R \ell_R \ell_c) \beta_{xT}^{(i)} \ddot{q}^{(i)} - \sum_{i=1}^n \frac{qS}{V} \sum_x \left[C_{(N/\alpha)_x} (x-x_{cg}) \phi_x^{(i)} \right] \dot{q}^{(i)} \\
& - \sum_{i=1}^n \left\{ T_c (\phi_{xT}^{(i)} + \ell_c \sigma_{xT}^{(i)}) + T_f (\phi_{xT_f}^{(i)} + \ell_c \sigma_{xT_f}^{(i)}) \right. \\
& \left. - qS \sum_x \left[C_{(N/\alpha)_x} (x-x_{cg}) \sigma_x^{(i)} \right] \right\} q^{(i)} - \sum_{\ell=1}^m [K_{\ell} \ell_{\ell s} + m'_{z\ell}] z_{\ell} \quad (187)
\end{aligned}$$

For rigid body plunging,

$$\begin{aligned}
M_0 \ddot{z} - \frac{F_{\alpha} \ell_{\alpha}}{V} \dot{\theta} + F_{\alpha} \alpha &= M_R \ell_R \ddot{\delta} + T_c \delta - \sum_{i=1}^n M_R \ell_R \beta_{xT}^{(i)} \ddot{q}^{(i)} \\
& + \sum_{i=1}^n \frac{qS}{V} \sum_x \left(C_{(N/\alpha)_x} \phi_x^{(i)} \right) \dot{q}^{(i)} - \sum_{i=1}^n \left[T_c \sigma_{xT}^{(i)} + T_f \sigma_{xT_f}^{(i)} \right. \\
& \left. + qS \sum_x \left(C_{(N/\alpha)_x} \sigma_x^{(i)} \right) \right] q^{(i)} + \sum_{\ell=1}^m K_{\ell} z_{\ell} \quad (188)
\end{aligned}$$

In these expressions the rigid body aerodynamic parameters are given by the following summations.

$$\begin{aligned}
 F_{\alpha} &= qS \sum_x C_{(N/\alpha)_x} \\
 F_{\alpha}^{\ell} &= qS \sum_x C_{(N/\alpha)_x} (x-x_{cg}) \\
 m_{\dot{\theta}} &= \frac{qS}{V} \sum_x C_{(N/\alpha)_x} (x-x_{cg})^2
 \end{aligned} \tag{189}$$

In summary, the equations of motion for the launch vehicle are given by Equations 174 (for angle of attack), 179 (for load torque opposing actuator motion), 184 (for slosh motion), 186 (for the flexible vehicle modes), and 187 and 188 (for the rigid body modes). In addition, an equation for the actuation torque is needed to complete the system of dynamics equations.

This set of equations is quite formidable. It would be desirable to simplify them, thus simplifying the analysis problem. For the class of medium sized, nonfinned launch vehicles considered in this example, it develops that the influence of aerodynamic forces, except on the rigid body degrees of freedom, is quite small. While it may be desirable to retain the aerodynamic damping terms on the various flexible modes, the influence of these forces in coupling one mode to the other, and in changing the modal frequency, is quite small for this class of vehicles. Even the aerodynamic damping can be neglected provided it is realized that the resulting solutions for closed-loop roots will be conservative (i. e., have less damping than in actuality). Further, the aerodynamic damping on the rigid body degrees of freedom can be neglected. The usual practice approximates α by θ for the intermediate frequencies, which removes consideration of Equation 174. A final consideration is to neglect the thrust-dependent terms that modify the flexible mode natural frequencies. The resulting equations are:

Rigid Body Pitching

$$\begin{aligned}
 \left(I_0 s^2 - F_{\alpha}^{\ell} \right) \theta &= \left[(I_R + M_R \ell_R \ell_c) s^2 + T_c \ell_c \right] \delta - \sum_{\ell=1}^m \left[K_{\ell} \ell_{\ell} + m'_{z\ell} \right] z_{\ell} \\
 &- \sum_{i=1}^n \left\{ (I_R + M_R \ell_R \ell_c) \beta_{xT}^{(i)} s^2 + \left[T_c (\phi_{xT}^{(i)} + \ell_c \sigma_{xT}^{(i)}) \right. \right. \\
 &\left. \left. + T_f (\phi_{xT_f}^{(i)} + \ell_c \sigma_{xT_f}^{(i)}) \right] \right\} q^{(i)}
 \end{aligned} \tag{190}$$

Rigid Body Plunging

$$M_0 s^2 z + F_\alpha \theta = \left(M_R \ell_R s^2 + T_c \right) \delta + \sum_{\ell=1}^m K_\ell z_\ell - \sum_{i=1}^n \left\{ M_R \ell_R \beta_{xT}^{(i)} s^2 + \left(T_c \sigma_{xT}^{(i)} + T_f \sigma_{xT_f}^{(i)} \right) \right\} q^{(i)} \quad (191)$$

Flexible Vehicle Bending (k^{th} mode)

$$\begin{aligned} & \left\{ \left[\mathfrak{M}^{(k)} - M_R \ell_R \beta_{xT}^{(k)} \phi_{xT}^{(k)} - \left(M_R \ell_R \phi_{xT}^{(k)} - I_R \sigma_{xE}^{(k)} \right) \beta_{xT}^{(k)} + I_R \beta_{xT}^{(k)} \sigma_{xT}^{(k)} \right] s^2 \right. \\ & \left. + 2 \zeta^{(k)} \omega^{(k)} \mathfrak{M}^{(k)} s + \mathfrak{M}^{(k)} (\omega^{(k)})^2 \right\} q^{(k)} \\ & = - \left[\left(M_R \ell_R \phi_{xT}^{(k)} - I_R \sigma_{xT}^{(k)} \right) s^2 + T_c \phi_{xT}^{(k)} \right] \delta - \sum_{\ell=1}^m \left[K_\ell \phi_{x\ell}^{(k)} + \mathfrak{M}'_{z\ell} \right] z_\ell \\ & + M_R \left[\phi_{xT}^{(k)} - \ell_R \sigma_{xT}^{(k)} \right] s^2 z + \left[M_R (\ell_R + \ell_c) \phi_{xT}^{(k)} - (I_R + M_R \ell_R \ell_c) \sigma_{xT}^{(k)} \right] s^2 \theta \\ & + \sum_{\substack{i=1 \\ i \neq k}}^n \left\{ \left[M_R \ell_R \beta_{xT}^{(k)} \phi_{xT}^{(i)} + \left(M_R \ell_R \phi_{xT}^{(k)} - I_R \sigma_{xE}^{(k)} \right) \beta_{xT}^{(i)} - I_R \beta_{xT}^{(k)} \sigma_{xT}^{(i)} \right] s^2 \right. \\ & \left. + \left[T_c \phi_{xT}^{(k)} \sigma_{xT}^{(i)} + T_f \phi_{xT_f}^{(k)} \sigma_{xT_f}^{(i)} \right] \right\} q^{(i)} \quad (192) \end{aligned}$$

This is about as far as the approximations can be taken without getting into specific cases of analyzing the stability of particular modes at particular flight times. In these instances, certain observations on the relative magnitudes of the various terms may permit sloshing, perhaps even the rigid body modes, to be neglected when analyzing for bending stability.

4.3.2 SOLUTIONS TO THE EQUATIONS OF MOTION. Solution of the flight control system/vehicle stability problem may be conveniently divided into two parts, the first broadly described as the frequency response approach, the second as the time response solution. In the former, the equations are not really solved. Only the significant parameters of the general solution are obtained; from these the necessary system design parameters are inferred to achieve the required degree of stability. The time response solution is obtained by means of a detailed simulation of the launch vehicle.

The simultaneous set of differential equations is solved and the solution measured against response criteria.

4.3.2.1 Frequency Response. The perturbational equations of motion of the flexible space launch vehicle consist of a large set of simultaneous ordinary differential equations. When combined with the equations describing the vehicle autopilot, the resulting set can be written in matrix form.

$$[A] \dot{x} = [b] \theta_c + [C] \alpha_w \quad (193)$$

where

$$\begin{aligned} [A] &= n \times n \text{ matrix of polynomial (in the Laplace transform variable, } s) \text{ coefficients} \\ x &= n \times 1 \text{ column vector of the dependent system variables} \\ [b] \text{ and } [C] &= n \times 1 \text{ column vectors of polynomial coefficients} \\ \theta_c \text{ and } \alpha_w &= \text{independent system variables (command and aerodynamic disturbance)} \end{aligned}$$

This matrix equation must be solved or otherwise manipulated to yield answers to the questions of vehicle frequency response and stability properties. Machine computation techniques are used to attain the required accuracy in a reasonable period of time.

4.3.2.1.1 Root Loci. The determinant of $[A]$ when set equal to zero is the characteristic equation of the overall system, the roots of which determine the system stability qualities. There exist many digital computer routines by which $[A]$ can be reduced to a polynomial in s from which the roots can be extracted. Other routines are available by which the roots are extracted directly without the intermediate step of obtaining the polynomial. The application to root locus studies is obvious. Consider the multiloop system of Figure 36. Two of the equations in the matrix Equation 193 are given by

$$\begin{aligned} D(s) \delta_c - K_A N(s) \theta_e &= 0 \\ \theta_e + K_1 \theta_{f_1} + K_2 \theta_{f_2} + K_3 \theta_{f_3} &= 0 \end{aligned} \quad (194)$$

The poles of the system under variation of the forward loop gain K_A are given by setting K_A equal to zero and solving the resulting characteristic equation. Repeated solution for various values of K_A yield the system roots which, when plotted on the complex s -plane, yield the root loci under variation of K_A . The zeros are obtained by setting θ_e in the second equation to zero with K_A at any finite value (equivalent to setting K_A equal to infinity). Similarly, the root loci under other parameter variations may be found. One can, for example, obtain repeated solutions of the characteristic

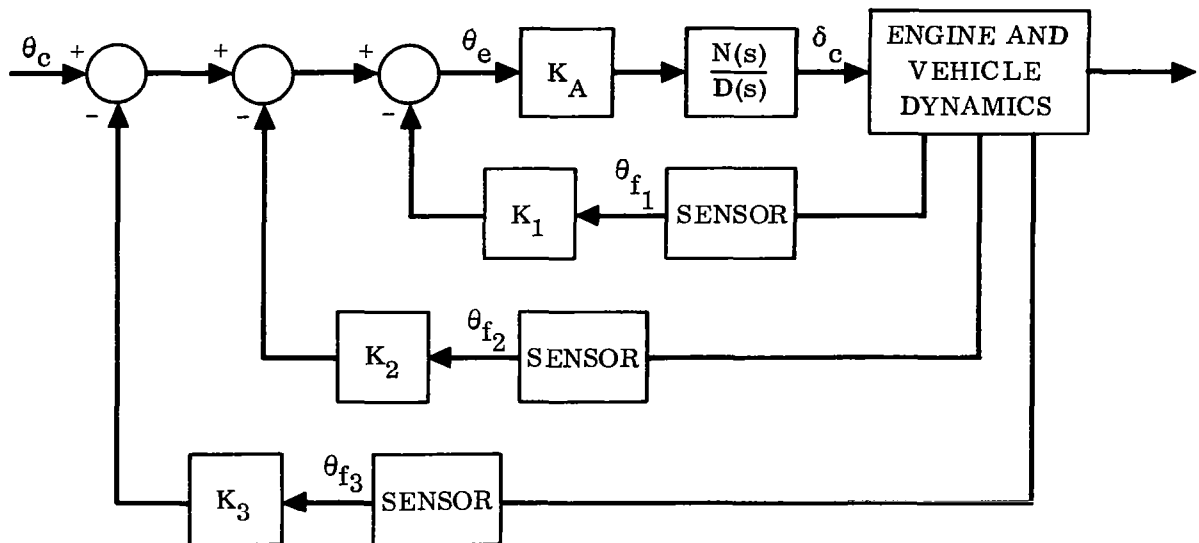


Figure 36. Block Diagram of Vehicle and Autopilot

equation as various system parameters are varied over their tolerance bands, thus yielding the root locations that will result for expected variations in these parameters. Inspection of the root locations reveals the stability qualities of the system.

The major limitation in this approach is brought about by systems having characteristic equations of high degree. Although there are a number of polynomial-solving routines available, each employing various iterative schemes for extracting roots, it cannot be guaranteed that at least one of them will converge for any particular case. If the polynomial coefficients are such that the roots are located in a saddle point region, some polynomial routines fail to converge; others converge (apparently) while giving erroneous answers. Other routines will give erroneous answers for roots lying on loci emanating from poles of small residue (a closely spaced and isolated pole-zero pair, for example). In these cases the roots in a saddle point region are very sensitive to small variations in system parameters, and the answer obtained must be regarded with a certain degree of skepticism. In the latter case, the root will contribute negligibly to the system response. There still remain some open questions: for example, how does one identify the fact that a root is in a saddle point region? Such questions have an important bearing on the overall stability qualities of the system under normally expected variations or uncertainties in system parameters.

In those cases where polynomial routines fail to provide satisfactory answers, one is forced to examine the matrix of polynomial coefficients to see if, within a restricted frequency range, the equations of motion can be satisfactorily represented by a smaller set. While this prevents one from obtaining the complete loci with any degree of accuracy, small root motion from the poles can be adequately handled. Which approximations to make is largely a matter of engineering judgment, with the physical understanding of the system being of paramount importance.

In some instances the approach outlined in Paragraph 4.2.5.3 is perfectly satisfactory for obtaining solutions where polynomial routines fail. There is no reason why the technique outlined there cannot be programmed for machine solution using an iterative scheme. The first solution is used as the point about which the Taylor's series is expanded for the second iteration, and so on.

4.3.2.1.2 Nyquist Criterion. If Equation 193 is rewritten in truncated form with the error signal the independent variable, the resulting set of equations is amenable to solving for the open-loop frequency response. One simply sets the Laplace transform variable $s = j\omega$ in the matrix equation for various values of ω and solves for the ratio of the feedback signal to the error signal $GH(j\omega)$ by means of Cramer's rule. Nyquist, Bode, or Nichol's plots can be drawn up from the resulting data. Thus

$$\begin{bmatrix} A' (j\omega) \\ \vdots \\ \theta_f \end{bmatrix} = b' (j\omega) \theta_e \quad (195)$$

where

$$\begin{bmatrix} A' (j\omega) \end{bmatrix} = n \times n \text{ matrix of polynomial coefficients evaluated at } s = j\omega$$

$$\begin{bmatrix} \vdots \\ \theta_f \end{bmatrix} = n \times 1 \text{ column vector of dependent variables containing the total feedback signal, } \theta_f$$

$$b' (j\omega) = n \times 1 \text{ column vector of polynomial coefficients evaluated at } s = j\omega$$

$$\theta_e = \text{independent (for the open-loop case) variable}$$

This approach has one major advantage over the root locus method in that one is not solving a polynomial but only obtaining the ratio of two determinants of constant coefficients for each value of ω . Nyquist plots for large systems can be obtained where the root locus approach would fail to yield an answer.

On the other hand, it is possible to be misled by the results of such an analysis because one is not necessarily sure of the number of open-loop poles in the right-half s -plane. If one never solves for the open-loop poles of the system, this doubt can

exist, especially in heavily crosscoupled systems. For example, it is possible to obtain open autopilot loop flexible mode poles in the right-half plane because of load torque feedback terms on the gimbaled engine motion. Open-loop poles in the right-half plane can also result from propellant slosh modes under certain conditions. One should investigate this possibility if a "lobe" of the Nyquist plot corresponding to a parasitic mode of response loops around in the counterclockwise direction with increasing frequency. The question then is whether or not the unstable pole results from an error in the input data, or exists in fact.

4.3.2.2 Time-Varying Simulation. A time-varying parameter simulation of the launch vehicle is used in the final stages of the launch vehicle flight control system analysis and synthesis procedure to provide quantitative information on the dynamic response properties of the launch vehicle. The equations of motion (in full, not just the perturbational form) are set up in a simulation of vehicle motion through the atmosphere. How detailed these equations should be (that is, which and how many degrees of freedom) is a function of the quality of answers desired and the frequency content of the expected disturbances due to the atmosphere (wind profile and gusts) and guidance commands (or pitch-over program). Generally speaking, if the frequency content of the disturbances is significant up to, say, two cycles per second, then all degrees of freedom having closed-loop frequencies in the band of zero to two cycles per second should be included. The only exceptions are stable modes of oscillation having small residues (pole almost canceled by zero); for example, low-density fluid slosh in small tanks. These modes are difficult to excite. Even when excited, the forces and motions produced are insignificant relative to the other degrees of freedom. Trajectory-dependent terms, although of low frequency, are included as they have a great bearing on the structural loads, the ultimate criterion of design. In addition the various cross-coupling terms between the control planes due to center-of-gravity offsets, cross-products of inertia (if significant), crosstalk, and the like are included, although the detail is a function of the particular vehicle configuration simulated. The object is to simulate those degrees of freedom likely to be excited during flight or having a bearing on the structural loads so that the simulation results are a close imitation of what might be expected in flight.

Such simulations are quite complex, requiring considerable time and effort devoted to engineering of the simulation itself. They can be set up on an analog, digital, or hybrid computer, the choice depending on cost effectiveness considerations and the intended application of the results. One must consider the relative premiums placed on overall cost, calendar time required for a solution, expected future use of the simulation after the immediate problem is solved, and anticipated requests for new information carrying a premium on rapid response time.

4.3.2.2.1 Analog Computer Simulations. The major virtue of the detailed analog computer simulation is its rapid solution time and "Rep. Op." (repetitive operation) capability. If the highest-frequency mode is two cycles per second, the simulation (depending upon the computer capabilities) can be set up to run faster than real time

by a factor of perhaps five to ten. A booster flight phase lasting three minutes can be run off in, say, 20 seconds with the results being observed and (to a degree anyway) assimilated on the spot. The engineer can then "tweak" the system gains or filter parameters and rerun the problem. This is what is meant by a "Rep. Op." capability: repeated runs with parameter variations and decision making between each run. As such, the detailed analog simulation becomes a powerful design tool for optimizing the flight control system configuration as a function of flight time.

The major failings of analog simulations are their relatively long setup and check-out times and a tendency for lack of repeatability in answers from one day (or even one run) to the next. One's success with these problems is therefore a strong function of the equipment available, the facility checkout and maintenance procedures, and the skill of the personnel responsible for setting up the simulation. In particular, procedures must be devised by which equipment malfunctions can be detected quickly, thus minimizing wasted solution time (invalidated because of a subsequent discovery of an equipment malfunction; a function generator loses a diode, for example). To be cost effective, a large analog computer simulation must be intended for many runs (solutions), such that the setup and checkout time can be amortized over the number of solutions obtained. In fact, a major problem may well be that of making sense out of the many yards of recorder traces that are produced in a short time.

4.3.2.2.2 Digital Computer Simulations. The major advantages of a digital computer simulation lie in its ability to provide answers within a relatively short time from the decision to set up such a simulation and the guarantee (at least until the computer manufacturer comes up with a system change requiring reprogramming) that the same answers will be obtained each time for a given set of conditions. Equipment malfunctions on the digital machine usually shut it down; unlike the analog, which runs on merrily producing erroneous results. Subsequent runs on a given simulation (program) require negligible setup time except for the changes in input data. Thus the digital simulation offers rapid response capability to "what if" problems. (Example: What if the payload weight is increased by 1000 pounds?) An answer to this type of question is a matter of coding some data cards and the turnaround time in the computer facility, usually within 24 hours. The same type of question asked of an analog simulation is only feasible if the simulation is already up and running.

The disadvantages of the digital computer when used in comprehensive simulations are concerned with the longer running time and the higher cost per "solution minute." The differential equations are solved in a stepwise fashion. The number of steps taken per unit time is a function of the highest frequency degree of freedom in the problem. The running time also increases with the number of degrees of freedom included. Various artifices, such as using longer compute cycles on the lower-frequency modes and shorter cycles on the higher-frequency modes, can materially reduce the running time. At best, the current state of the art on the highest-speed digital machines only approaches real time, and is usually slower. Here again the skill of the programmer becomes important in reducing the cost of operation. Even so, large numbers of

digital computer runs will generally cost more than the same runs on the analog computer. The breakeven point is a function of the relative capabilities of the analog and digital machines available, and their respective costs of operation.

The longer running time and the need to utilize the expensive digital computer efficiently results in serious disadvantage with regard to "Rep. Op." relative to the analog computer. It is usually unfeasible to use the digital computer as a synthesis tool for optimization of the flight control system. The analog computer operator can make five iterations in as many minutes. The same job on the digital machine requires, say, two or three days for the iterative process, or more than five separate runs from which the best solution is chosen; either of which is more costly from the standpoints of time and money. This presumes, of course, that the digital computer program is equipped with a plotting option for visualizing the results.

The choice between simulation techniques is seen to be a strong function of the complexity of the problem being solved, the computers available, their operating costs, and the organizational structure (can the engineer get his job run now, or must he wait until the job comes to the head of the line in a batch-processing scheme?). The respective advantages of each approach has led to the development of the hybrid computer in an effort to get the best of both worlds.

4.3.2.2.3 Hybrid Computer Simulations. The hybrid computer is a relatively recent development that combines a special purpose digital computer with an automated analog computer. Software (digital computer program) items form an integral part of the system by which the analog portion is checked out, the functions and potentiometers set, and check solutions run; thus automating, to a large extent, the setup and checkout portions of the operation. In addition the digital computer works "on line" with the analog. In the type of problem under discussion the digital portion can simulate a guidance system and provide wind profile data while processing the dependent variables to produce flight loading information. The analog portion simulates the vehicle dynamics and those other high-speed parts of the program. The hybrid machine thus has the advantages of both digital and analog schemes; it also has the disadvantages of both. The former may be summarized as follows:

1. High Speed - integrations and high-speed arithmetic are performed on the analog; low-speed arithmetic and data handling on the digital.
2. Capability of simulating combined continuous/discrete systems (computer-controlled launch vehicles for example).
3. "Rep. Op." Capability - the digital machine is subject only to the demands of the analog; other uses of the digital processor are on a noninterference basis.
4. Automated setup and checkout of the analog portion of the machine.

The disadvantages may be similarly enumerated:

1. Repeatability - this analog computer bugaboo is still present although self-checking schemes can be built into the digital side of the system.
2. Cost - the initial setup, programming, and debugging times are longer than would be experienced for either analog or digital computers. And the digital portion may suffer from poor utilization. These factors plus the higher initial cost of the hybrid machine make operating costs per hour intermediate between the pure analog or pure digital machines. (By the same token, the results should be more general than pure analog results, faster than pure digital results, and in fact close to optimal as far as complex simulations are concerned. Subsequent rapid and easy use is a major benefit from the higher cost and may turn the cost factor into an advantage, but this has yet to be demonstrated.)
3. Complexity - the operator must combine the talents of digital computer programmer, analog computer operator, and the engineer who formulates the problem. The engineer's confidence in the answers obtained increases in direct proportion to his proximity and involvement in the solution of the problem. From the engineer's standpoint, then, it is he who must be capable of operating the machine. The alternative is to fall back to the closed shop concept used in some analog computing facilities and almost all digital computing facilities, while paying the inevitable penalty in turnaround time and general suspicion of the answers arrived at: the hybrid computer becomes surrounded with the mystique of the large digital data processor.

Figure 37 gives some indication of the comparative merits of the analog, digital, and hybrid computers in solving the time-varying simulation problem. The chart is admittedly and necessarily subjective and reflects a lower experience level with hybrid machines. Mechanization is felt to take longer, not only because of combined digital-analog operation, but also because of the choices that must be made as to what portion of the problem is best solved on one portion or the other. On the other hand, checkout time should be shorter on the hybrid machine than the pure analog because of its automated nature.

To summarize, it is felt that the digital computer solution to the simulation problem is most cost effective where small numbers of runs, fast response time, and minimal approximations in the equations of motion are contemplated. Hybrid (or analog) operation shows cost advantages when large numbers of runs are required (for example, a statistical analysis of inflight loads in response to a large number of wind profiles).

The only open question is the relative cost effectiveness of the cheaper analog computer versus the more expensive hybrid machine. The versatility of the latter is not usually needed for the class of problems under discussion here. In particular, for the flight loads problem the analog computer alone will suffice given a D-to-A, A-to-D

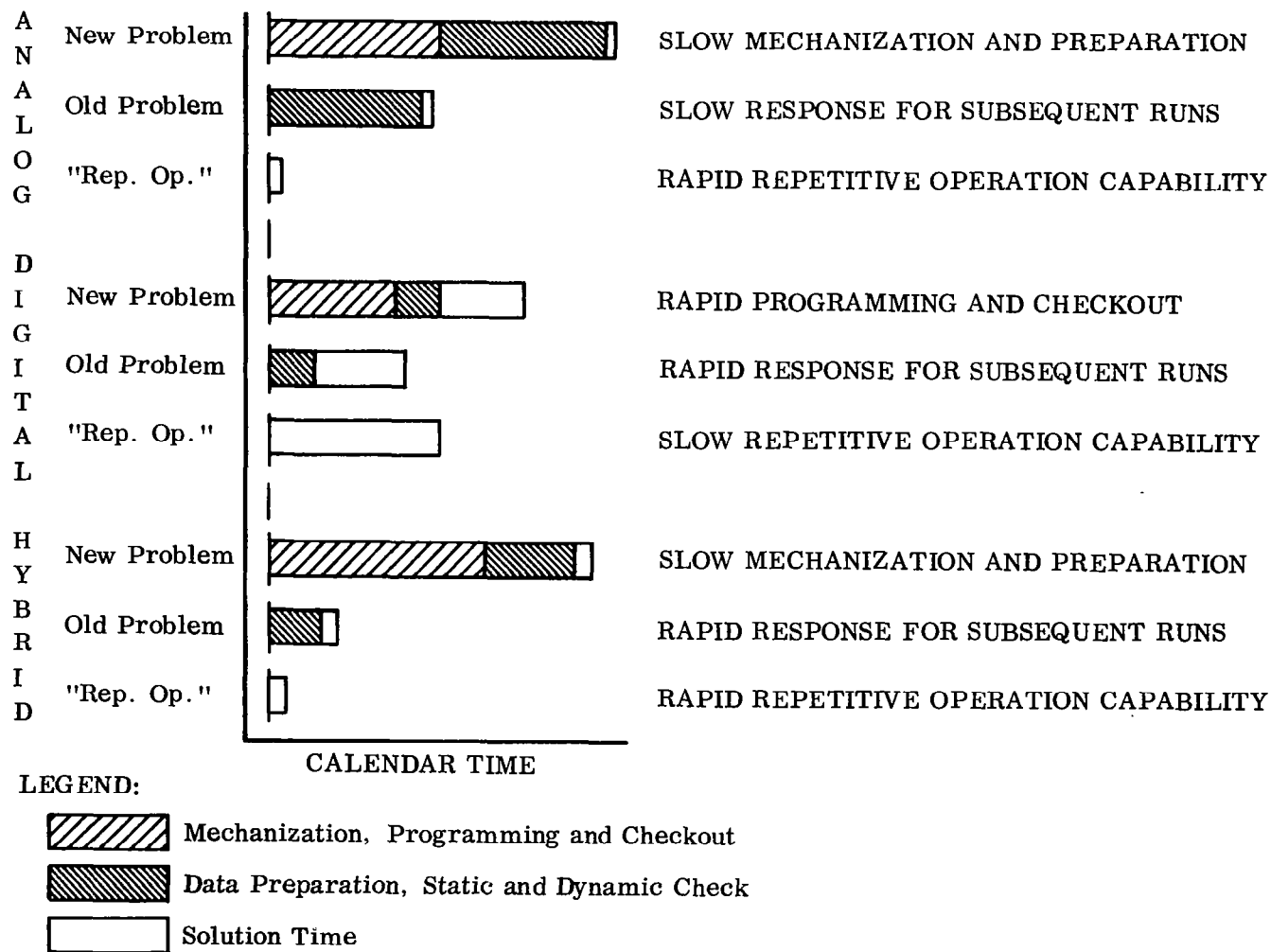


Figure 37. Comparative Solution Times

converter capable of translating recorded (on magnetic tape, for example) digital information to analog, and recording analog information in digital form. The data processing machine can be used to process the recorded information off-line. There are, of course, some instances where only the hybrid machine capabilities will suffice. The closed-loop guidance problem is a good example.

4.3.2.2.4 Flight Loads. The major criteria applied to the time response properties of a launch vehicle are the loading conditions to be expected in flight. The flight control system parameters, even apart from drift minimization or load reduction schemes, greatly influence the inflight loads because of their influence on the time response properties of the launch vehicle in the atmospheric environment. One major purpose of setting up a time-varying simulation of a launch vehicle is to obtain a measurement of the loads encountered as the simulated vehicle proceeds from launch through the atmosphere to exit.

The bending moment at station x_n is given by the summation of all moments, working back from the nose of the vehicle.

$$\begin{aligned}
 (BM)_{x_n} = \sum_j \left\{ -M_{x_j} \left(\sum_i \phi_{x_j}^{(i)} \ddot{q}^{(i)} + (x_j - x_{cg}) \ddot{\theta} - \ddot{z} \right) \right. \\
 \left. + (qS) C_{(N/\alpha)_{x_j}} \alpha_{x_j} \right\} (x_j - x_n) \\
 - \sum_\ell \left[K_\ell (x_{\ell s} - x_n) + m'_{z_\ell} \right] z_\ell
 \end{aligned} \tag{196}$$

where

$$\begin{aligned}
 M_{x_j} &= \text{lumped equivalent mass at station } x_j \\
 \alpha_{x_j} &= \text{defined by Equation 174} \\
 j &= \text{station index forward of } x_n
 \end{aligned}$$

and the summation of slosh moments is understood to refer only to those spring-mass and slosh moments forward of station x_n .

4.3.3 STATISTICAL INFLUENCES ON SYSTEM BEHAVIOR. The design verification of the flight control system cannot be considered complete until the full range of possible system behavior has been defined. This includes the variations in dynamic behavior due not only to the expected range of vehicle and flight control system parameters, but also to the range of possible flight environments. Investigation of this problem is done in varying levels of sophistication depending upon launch vehicle and

mission requirements. This paragraph briefly discusses these techniques and their application to the dynamic stability and launch availability problems.

4.3.3.1 Analytical Techniques. In increasing levels of precision, the techniques used to define the range of system performance range from the relatively crude but easy to apply "worst case" approach to the most sophisticated Monte Carlo approach. These are discussed below together with their range of application.

4.3.3.1.1 "Worst Case" Approach. The easiest means by which some idea of the range of system performance may be estimated is to assume that each variable influencing a particular performance parameter is at its worst possible value such that the worst effect is produced. For example the rigid body control mode damping at maximum dynamic pressure is a function of the autopilot rate, position and integral gains, the control actuation and autopilot filter lags, and the control moment effectiveness and aerodynamic instability parameters μ_δ and μ_α (the latter is a function of the trajectory flown). If the autopilot rate and position gains and μ_δ are at their minimum values while all other parameters are at their maximum values, the rigid body control mode damping will be at the minimum value it can attain.

This approach, while straightforward, is highly conservative; so much so as to be unrealistic. It is unlikely that all autopilot parameters on a particular vehicle will assume their worst values. It is such a conservative approach that in many cases one cannot afford to use it. In this example, its application could easily result in autopilot tolerances considerably tighter than could reasonably be guaranteed for reasonable cost. On the other hand, where it is possible to use the worst-case approach is where the predicted extremes of behavior have little effect on the overall system stability and performance qualities. This is where it offers an easy means of evaluation, which maximizes the confidence level in system design. The extremes of system behavior, besides having little influence on overall qualities, also have little if any statistical significance except to say that they are highly improbable.

4.3.3.1.2 Statistical Formulation. If there exists (or can be assumed) some statistical knowledge of the variables that determine a performance measure, one can treat the latter as a statistical quantity, thereby assigning a probability density function to it. This allows one to make statements concerning the relative probability of various values of the performance measure.

The crudest approach is the well-known root-sum-squared (rss) technique. Consider a dependent random variable w that is a function of n statistically independent random variables x_i ($i = 1, 2, \dots, n$):

$$w = \phi(x_1, x_2, \dots, x_n) \quad (197)$$

Each of the x_i has a probability density function $p_i(x_i)$ associated with it. This function defines a mean \bar{x}_i and a standard deviation σ_{x_i} of the random variable x_i . Note that $p_i(x_i)$ can assume any form (x_i does not have to be a normally distributed random variable, only statistically independent of all other x_i). The following relations can be shown:

$$\bar{w} = \bar{\phi} + \frac{1}{2} \sum_{i=1}^n \frac{\overline{\partial^2 \phi}}{\partial x_i^2} \sigma_{x_i}^2 + \dots \quad (198)$$

$$\begin{aligned} \sigma_w^2 = & \sum_{i=1}^n \left[\left(\frac{\overline{\partial \phi}}{\partial x_i} \right)^2 - \frac{1}{4} \left(\frac{\overline{\partial^2 \phi}}{\partial x_i^2} \right) \sigma_{x_i}^2 \right] \sigma_{x_i}^2 \\ & + \frac{1}{2} \sum_{\substack{i=1 \\ j=1 \\ i \neq j}}^n \left[\left(\frac{\overline{\partial^2 \phi}}{\partial x_i \partial x_j} \right)^2 + \frac{\overline{\partial \phi}}{\partial x_i} \frac{\overline{\partial^3 \phi}}{\partial x_i \partial x_j^2} + \frac{\overline{\partial \phi}}{\partial x_j} \frac{\overline{\partial^3 \phi}}{\partial x_i^2 \partial x_j} \right. \\ & \left. + \frac{1}{2} \bar{\phi} \frac{\overline{\partial^4 \phi}}{\partial x_i^2 \partial x_j^2} \right] \sigma_{x_i}^2 \sigma_{x_j}^2 + \dots \end{aligned} \quad (199)$$

where the barred quantities indicate evaluation at \bar{x}_i ($i = 1, 2, \dots, n$)

$$\begin{aligned} \bar{w} &= \phi(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) \\ \sigma_w^2 &= \sum_{i=1}^n \left[\frac{\overline{\partial \phi(x_1, x_2, \dots, x_n)}}{\partial x_i} \right]^2 \sigma_{x_i}^2 \end{aligned} \quad (200)$$

If there is any question regarding the validity of the linear approximations, one can calculate the next higher-order derivatives to check their contribution to the total summation.

This formulation tells us that the variance of w is a linear function of the variances of the x_i , without making any assumption as to the form of the probability density function for either the x_i or w . Chebyshev's inequality* holds, which allows one to draw limited conclusions concerning the probability of various values of w .

*Chebyshev's inequality states that $P(|w - \bar{w}| \geq \epsilon) \leq \frac{\sigma_w^2}{\epsilon^2}$

The central limit theorem allows one to draw further conclusions concerning the probability density function for w . For example, if the x_i are normally distributed random variables, then so is w . Further, if there are many non-normally distributed x_i having approximately equal contributions to w , then $p(w)$ tends toward a normal distribution as the number of the x_i becomes large. Thirdly, if the dominant x_i influencing w are normally distributed, $p(w)$ will tend toward a normal distribution as the influence of the remaining non-normally distributed x_i becomes small, or their number becomes large.

In some cases one has a limited sample of possible values for the x_i . If one assumes that the x_i are normally distributed random variables, then one can specify both mean and variance on x_i , but with a certain confidence level. The ultimate result is to associate mean and variance with a level of confidence; the variance increases as the confidence level increases and vice versa. A discussion of this topic can be found in many textbooks on probability and statistics; for example, Reference 10.

In those cases where the form of the probability density function $p(w)$ for the random variable w (the performance measure or parameter in this discussion) must be determined in order to make sufficiently precise statements concerning the probability of various values of w , one must resort to a more complex mathematical procedure. Knowing the functional form of $\phi(x_1, x_2, \dots, x_n)$ and the probability density functions $p_i(x_i)$ ($i = 1, 2, \dots, n$), the form of $p(w)$ can at least in theory be determined. For example, if $w = \sum_{i=1}^n x_i$ then $p(w)$ is equal to the successive convolution of the individual $p_i(x_i)$ (Reference 11). If one employs characteristic functions defined by

$$M_{x_i}(j\nu_i) = \int_{-\infty}^{\infty} e^{j\nu_i x_i} p_i(x_i) dx_i \quad (201)$$

then it can be shown that

$$M_w(j\nu) = \prod_{i=1}^n M_{x_i}(j\nu) \quad (202)$$

whence

$$p(w) = \frac{1}{2\pi} \int_{-\infty}^{\infty} M_w(j\nu) e^{-j\nu w} d\nu \quad (203)$$

In practice, $\phi(x_1, x_2, \dots, x_n)$ is usually too complex to readily yield to this approach, and one is faced with two alternatives: either be content with the more limited interpretations possible with root-sum-squaring or use the "brute force" approach: solution by Monte Carlo techniques.

4.3.3.1.3 Monte Carlo Simulations. The Monte Carlo approach consists of what might be termed an experimental solution to obtain $p(w)$ and various properties (mean, standard derivation, limits on w that will not be exceeded for a specified percentage of the possible values of w , etc.) of the probability density function. The only requirement is that the $p_i(x_i)$ be known or suitably approximated. The functional form of $\phi(x_1, x_2, \dots, x_n)$ need not be known explicitly although it is implicit in the simulation. The procedure is as follows:

1. Set up a simulation by which w can be determined given a set (x_1, x_2, \dots, x_n) . This may consist of a simple mathematical expression, but can also be, say, a simulation of a launch vehicle flight with provision for determining the in-flight loads.
2. Using one or another means for generating random numbers, pick a random set of independent variables (x_1, x_2, \dots, x_n) that satisfy the individual $p_i(x_i)$ and run the simulation to obtain a sample value for w .
3. Repeat this process for different selections of the set of independent random variables a sufficient number of times that the required level of confidence in the results is obtained.
4. Analyze, using standard statistical techniques, the results of these many repeated runs to obtain the statistical properties of w .

This approach can be used in almost any situation where the time and cost of running a Monte Carlo simulation are justified by the requirements for answers having a high confidence level. There are almost no restrictions on its range of applicability. The independent variables do not even have to be statistically independent, although this implies a somewhat more sophisticated means for picking the set (x_1, x_2, \dots, x_n) . In fact, this approach to the solution of the statistical problem is justified only in very complex situations where there are many factors that do not readily yield to techniques other than direct simulation. The key to the success of the procedure is in picking the sets (x_1, x_2, \dots, x_n) in the required fashion. The results obtained are a direct function of how closely these sets approximate the desired conditions of statistical dependence and independence, as well as satisfying the requirements of $p_i(x_i)$.

4.3.3.2 Flight Control System Parameter Tolerances. One of the analytical tasks facing the control engineer is that of defining the range of system behavior expected as a result of expected variations in the parameters defining the flight control system. The synthesis task is the same problem in reverse; defining the required tolerances on system parameters which will result in the system performance falling within an acceptable range. The analytical task is straightforward, using the techniques of the preceding paragraph; however, the same cannot be said for the synthesis task. The problem is to define the range of acceptable performance, a task requiring application of engineering judgment in most cases although in some it can be rigorously defined from the circumstances of the situation.

For example, consider the stability properties of the first bending mode throughout the booster phase of flight. Due to uncertainties (purely analytical) in the mathematical model it is more or less arbitrarily decided that the stability of this system root shall be such that a gain margin of ± 6 db and a phase margin of ± 30 deg is required throughout flight, regardless of the range of possible parameter variations that might be encountered within the flight control system. If the nominal phase and gain characteristics of the flight control system are such that the phase lag at first-mode frequencies increases by, say, 80 degrees between launch and the end of booster phase due to the first-mode frequency change, one is left with an allowable tolerance on flight control system lag of +20 deg at liftoff and -20 deg at the end of booster phase at these respective frequencies. If the system design is such that the hardware can be guaranteed to these tolerances, the problem is solved. If not, two alternatives arise, first to reconfigure the system such that the problem does not occur, second to reduce the phase margin requirements.

The foregoing is an example of the tradeoff between control system allowable tolerances and performance requirements. Stringent requirements in the area of response and stability margins imply restrictive parameter tolerances. The control engineer's problem is to simplify the control system configuration to the greatest extent possible with the widest tolerances on system parameters while retaining acceptable performance measures. In cases where the situation is very "tight"; that is, considerable difficulty in obtaining system components guaranteed to meet the imposed system tolerances; one may be forced into "fine tuning" procedures where components are selectively assembled to meet system specifications. This imposes considerable difficulty in the field when components may have to be replaced. A second difficulty is in the testing procedure itself; a portion of the tolerance band is imposed by the test equipment accuracy of measurement. In short, the range of acceptable system performance, while perhaps arbitrary to some extent, must be chosen with care in order to avoid unnecessary difficulties later on.

4.3.3.3 Launch Probability and Availability. In large launch vehicles the probability of a successful launch may be heavily dependent on the winds aloft. To put it another way, the availability of the launch vehicle for a particular launch window is dependent on the system configuration (load-relieving features, for example) and load-sustaining capability. The launch probability or launch availability is therefore a fundamental performance measure of the launch vehicle.

Evaluating vehicle capability in this respect is not as straightforward as in others. The major problem is that of deciding what constitutes a statistically likely atmospheric environment. Historically, there have been two approaches. The first is to use an artificial (synthetic) wind profile so constructed as to yield the worst expected loading condition for which the vehicle is to be designed. Unfortunately, the worst condition for one vehicle may not be for the next, which implies that the wind profile used is a function of the launch vehicle configuration. The second approach is to use a statistically significant sample of measured wind profiles and evaluate launch

availability by a Monte Carlo technique where the independent variable is the wind profile used. Both approaches require superposition of artificial (as opposed to measured) wind gusts to obtain the complete loading condition.

4.4 CORRELATION OF ANALYTICAL PREDICTIONS AND TEST RESULTS

A major part of any launch vehicle development effort is the associated test program wherein the various vehicle systems, subsystems, and components are checked out for compliance with design intent and systems compatibility. Relative to the control analyst's task in the development effort, such testing constitutes his major source of information by which he is enabled to predict the dynamic behavior of flight. The analyst must necessarily make approximations to facilitate solution of his problems. The test program serves to check out the adequacy of his mathematical models.

Testing takes place in three phases. The first is what may be termed model development testing wherein the dynamic parameters of the launch vehicle are experimentally determined. Preflight and production testing of components and system is second. The purpose here is to ensure that the hardware meets system requirements and to ascertain the dynamic properties of the system as flown, enabling meaningful evaluation of postflight test results. Flight testing is the third phase and is the ultimate proof of the adequacy of a launch vehicle design. The last two of these categories of testing are discussed below with regard to the interpretation of test results; model development testing is discussed in another volume in this series of monographs.

4.4.1 PRODUCTION TESTING AND PREFLIGHT CHECKOUT. Preflight and production testing is basically a quality control function, the major purpose being to assure compliance with the design requirements for the various components and systems on the launch vehicle. The testing philosophy used is that which maximizes cost effectiveness. Large production runs of weapons system launch vehicles have production testing geared to a go/no-go basis, the idea being to minimize the cost per launch while accepting some percentage failure rate. Launch vehicles have a different emphasis. Here the desire is for maximum reliability on a per launch basis, with a higher cost per launch being tolerated. In the former case, one can make statistical judgments of probable system performance if one has access to the statistics of the component and systems variables. In the latter, the total number is small enough such that statistics are less meaningful and each vehicle is treated as a separate case. In any launch vehicle program the testing philosophy will lie somewhere between the two extremes, and in fact will be analogous to the design philosophy (see Paragraph 4.2.8). Generally speaking, the more "highly tuned" a launch vehicle is, the more stringent are the testing requirements, component tolerances, and the like; and the more important the results of such testing are to the control engineer.

In the former case, the control analyst has specified limits on the parameters of such hardware components as gyros, filters, electrohydraulic actuators, and the like.

But without knowing the statistics of such parameters, he cannot predict what constitutes "nominal" or "three-sigma" behavior. For this reason statistical information on subsystem components is necessary. In particular, both the nominal values and the variances on component parameters must be known. There is the further obvious requirement that the parameters measured be meaningful to the analysis. Systems tests and preflight checkout are oriented to giving a go/no-go answer, and do not usually satisfy this requirement.

The latter case is more pertinent to the development program (weapons system or otherwise) and those programs having relatively small production. Here the statistical nature of the component parameters may not yet be established (insufficient experience) or the number of components of a particular type is too small to permit valid statistical inference. Since the experience level is so low (each vehicle is a new "bird") the control engineer must know as much as possible about the vehicle systems prior to launch. Differences between the predicted and expected behavior can then be ascribed to unknowns, chiefly vehicle dynamic properties. The testing philosophy is such that each component, subsystem, and system is tested not only to see that it is working properly and falling within limits set up by the system requirements, but also to ascertain the actual parameters pertinent to the particular launch. The requirements for testing precision, checkout time, and documentation are correspondingly greater, the price that must be paid for maximum assurance of launch success. The control engineer's requirements for data and precision of measurement are basically those physical quantities that correspond to the numbers and mathematical models used in his analysis. Of paramount importance are the key parameters of flight control system gain and frequency response. From these data the inflight behavior can be predicted.

4.4.2 FLIGHT TESTING. The final proof of the design adequacy of a launch vehicle rests with the flight tests. Unfortunately, if the control engineer has done his job well, there is little quantitative information on stability margins to be obtained. However, careful evaluation of flight test results can go a long way toward verifying the mathematical models used in the analysis. The data acquired are limited by the instrumentation available and by the fact that only a few parasitic modes may be excited, and these only under transient conditions. Unstable modes or modes with some limit cycle character will be observed at other times. In either case one has a check on the frequency and damping ratio (if stable) or amplitude (if limit cycle).

Low-frequency modes of oscillation (rigid body, propellant sloshing, and low-frequency bending modes) will be excited during passage through the atmosphere by imposed wind gusts near the natural frequencies of these modes. Excitation by the guidance system or the pitch program may also provide information if such excitation is within the proper frequency range. The higher-frequency modes are only excited by the higher-frequency transients: launch, burnout, staging, etc. In these instances the frequencies may be checked, but the damping ratios (corresponding to closed-loop roots) may easily differ from prediction by factors of two or three. They are not a

good check unless gross analytical errors or omissions have been made. Successful postflight evaluation rests on precise knowledge of the flight control system parameters as flown. If a rigid body limit cycle is observed at a larger than expected amplitude and at lower frequency, one can attribute this to a larger than expected value of the aerodynamic instability parameter or to rate and/or position gains on the low side of the tolerance band. If the latter are known, an estimate of the former is possible; if not, then no conclusion can be drawn.

4.5 SUMMARY OF RECOMMENDED PRACTICE

This section has presented the techniques and methods finding greatest use in the analysis and synthesis of launch vehicle flight control systems. The format has been oriented toward providing an orderly transition between standard textbook methods and their application to the launch vehicle stability problem. The major difference between the two is one of complexity: launch vehicles are characterized by many parasitic modes of response that must be allowed for in control system design. While the use of modern computing machines is indicated for the solution of these problems, there exist many cases where the vehicle dynamics can be approximated over a restricted range of frequencies, making hand computation techniques feasible. There is an unfortunate tendency on the part of control engineers to rely too extensively on the computational capabilities of the machine, by which valuable insight to the essentials of the problem can be lost. With this in mind, the greatest attention has been paid to those approximations and methods amenable to hand calculation for use in the preliminary design phases of a launch vehicle program. The more complex situation follows as an extension.

The two most powerful techniques in frequency domain analysis, the root locus and frequency response methods, have been presented in such a way as to show the close relationship between the two. These techniques have been well developed in the control literature. In this monograph the emphasis has been on the newer approach as being less well known among control analysts and therefore more instructive. Also, root locus plot is somewhat easier to interpret in terms of the relative stability of the system roots under parameter variations, and the consequent influence on system response. However, the older approach offers computational advantages when analyzing systems of high order; one does not have to factor high-order polynomials in machine solutions. The question of stability margins has been handled in a somewhat more rigorous fashion than a more or less arbitrary selection of gain and phase criteria. The stability margin used is a function of the particular mode being considered, the analytical and statistical uncertainties associated with the mode, and response requirements. No general statement of stability criteria can be given, only the means by which reasonable criteria can be selected. This process still requires exercise of judgment in the area of analytical uncertainties in the mathematical model, but the statistical uncertainties (from all sources) and the response requirements are subject to rigorous evaluation.

Approximations to the vehicle dynamics have been made on the basis of frequency separation as well as secondary effects. For example, a frequency separation of two to one can usually permit neglecting the higher-frequency mode when studying the stability of the lower, and vice versa. Even when the frequencies are closer together, one can still draw conclusions on the stability qualities of these modes; i.e., which parameters improve or deteriorate stability, even though the actual numbers arrived at are in error. Secondary effects that change certain parameters by a few percent can also be neglected until detailed studies justify their inclusion.

The subject of nonlinearity in vehicle behavior has been treated briefly, in keeping with the intent of the monograph to restrict attention to linear analysis. Such nonlinearities that do exist result from amplitude-dependent slosh damping when using baffles and from electrohydraulic (or pneumatic) thrust chamber positioning servos. In both these instances the need for ground testing or reliable empirical data is indicated to verify the adequacy of the mathematical model.

The success of the analyst in configuring the flight control system for a launch vehicle is a direct function of the accuracy of the mathematical models used in all areas. Future effort should be concentrated along these lines: analytical, experimental (ground and flight testing), and correlation between the two. Most "follow on" work in the course of a launch vehicle program is necessarily concentrated on postflight analysis to improve the mathematical models and to gain experience where theory is inadequate.

4.6 REFERENCES

1. D. McRuer, Methods of Analysis and Synthesis of Piloted Aircraft Control Systems, BuAer Report No. AE-61-41.

This report includes a discussion of many derived performance criteria as well as a tabulation of some of the standard forms.

2. J. G. Truxal, Automatic Feedback Control System Synthesis, McGraw-Hill, New York, 1955.

This text on frequency domain analysis and synthesis is probably the most popular with control engineers for its clarity of exposition.

3. C. H. Wilts, Principles of Feedback Control, Addison-Wesley, Reading, Mass., 1960.

This newer text is more rigorous than the preceding where the exposition is more closely tied to the differential equations of motion of the system. Much of the discussion of the root locus and frequency response methods in this monograph is taken from this source.

4. C. F. Chen, "A New Method for Finding Breakaway Points of Root Loci Involving Complex Roots," IEEE Trans. on Auto. Control, pp 373, 374, July 1965.
5. R. V. Churchill, Introduction to Complex Variables and Applications, McGraw-Hill, New York, 1948.

Churchill's text is perhaps the most popular one on complex variable theory and the concept of mapping.

6. E. Polak, "A Note on D-Decomposition Theory," IEEE Trans. on Auto. Control, pp 107 - 109, January 1964.
7. A. L. Greensite, Analysis of Liquid Propellant Mode Stability of a Multi-tank Booster Vehicle, JAS, Vol. 29, #2, pp 130 to 139, February 1962.

This paper discusses the applicability of the root locus technique in factoring the numerator and denominator polynomials associated with vehicle transfer functions including sloshing. (It has been extended somewhat in this monograph, Paragraph 4.2.6.1, and can be applied to the flexible vehicle case, at least in principle.) Each additional mode is treated as another closure of a feedback loop.

8. D. R. Lukens, A. F. Schmitt, and G. T. Broucek, Approximate Transfer Functions for Flexible-Booster-and-Autopilot Analysis, WADD Report No. TR-61-93, April 1961.

This report covers much of the same ground as this monograph and includes a derivation of the describing function for hydraulic actuators as well as a discussion of the details involved in modeling a flexible launch vehicle. A discussion of component (roll gyros, accelerometers, etc.) transfer functions is also included.

9. R. L. Bisplinghoff, H. Ashley, and R. L. Halfman, Aeroelasticity, Addison-Wesley, Reading, Mass., 1955.

The classic text on the subject.

10. I. W. Burr, Engineering Statistics and Quality Control, McGraw-Hill, New York, 1953.

This text emphasizes the experimental approach to statistics.

11. W. B. Davenport, Jr., and W. L. Root, An Introduction to the Theory of Random Signals and Noise, McGraw-Hill, New York, 1955.

This text presents a rigorous approach to probability and statistics before going on to the subject matter of the title.

010 001 57 51 3DS 68059 00903
AIR FORCE WEAPONS LABORATORY/AFWL/
KIRTLAND AIR FORCE BASE, NEW MEXICO 87117

ATT MISS MADELINE F. CANOVA, CHIEF TECHNICAL
LIBRARY /WLIL/

POSTMASTER: If Undeliverable (Section 158
Postal Manual) Do Not Return

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Notes, and Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION DIVISION
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546